Planetary and Space Science

Volume 2 1959

Editor-in-Chief: MILTON GREENBERG

Regional Editors:

D. R. BATES
W. W. KELLOGG
A. I. LEBEDINSKY
D. F. MARTYN
S. K. MITRA
M. NICOLET
F. L. WHIPPLE

PERGAMON PRESS

NEW YORK · OXFORD · LONDON · PARIS

Planetary and Space Science

NOTES FOR CONTRIBUTORS

1. Only papers not previously published will be accepted, and emphasis is placed upon originality of concept or observation. The text should be clear and concise, and a technical description of the methods used should be given in detail only when the methods are new.

2. The Editor-in-Chief requests that short *Research Notes* for which rapid publication is desired should be sent directly to him at Geophysics Corporation of America, 700 Commonwealth Avenue, Boston 15, Massachusetts, U.S.A. *Research Notes should be in English, without illustrations, and no longer than four double-spaced typewritten pages.*

3. Other contributions (Research Papers and Critical Surveys) should be sent initially to one of the other editors near the author: in the *Southern Hemisphere*, to Dr. D. F. Martyn, Radio Research Laboratories, CSIRO, Camden, New South Wales, Australia; in the *Far East*, to Prof. S. K. Mitra, F.R.S., Institute of Radiophysics and Electronics, University of Calcutta, India; Prof. T. Nagata, Geophysical Institute, Tokyo University, Tokyo, Japan; in the *United Kingdom*, to Professor D. R. Bates, Department of Mathematics, Queen's University, Belfast, Northern Ireland; in *Europe*, to Dr. M. Nicolet, c/o Service du Rayonnement, 3 Avenue Circulaire, Uccle, Belgium; in *Eastern North America*, to Dr. Fred L. Whipple, Astrophysical Observatory, 60 Garden Street, Cambridge 38, Massachusetts; in *Western North America*, to Dr. W. W. Kellogg, Head, Planetary Sciences, The Rand Corporation, 1700 Main Street, Santa Monica, California, U.S.A.; and in the *Soviet Union*, to Prof. A. I. Lebedinsky, Physical Department, Moscow State University, Moscow, U.S.S.R.

4. Submission of two typescripts (an original and a carbon copy) will facilitate refereeing; if one only is submitted, it should be the original. The author should retain a copy for his own use. The text should be ready for setting in type and should be carefully checked for errors. Scripts should be typed on one side of the paper with double or triple spacing.

5. Illustrations may accompany Research Papers and Critical Surveys but should be restricted to the minimum necessary, and must accompany the typescript mounted on separate sheets. Photographs should be enlarged sufficiently to permit clear reproduction (about 8 by 10 inches). If words or numbers appear in photographs, two copies should be sent, one with clear printing and the other without inscription. Line drawings which require redrawing should include all relevant details. Any lettering should be large and bold enough to be legible after the diagram has been reduced in size for printing. Captions or legends should be given on a separate sheet, and labelled to show which illustrations each is to accompany.

6. Proofs will be sent to authors for correction, and fifty free reprints of each paper will be provided. Additional reprints should be ordered when the proofs are returned by the author. Otherwise, it may not be possible to furnish them at a reasonable price. A reprint order form will accompany the proofs.

7. References are indicated in the text by superior numbers in parentheses, and the full reference should be given in a list at the end of the paper in the following form:

1. Y. TANAKA, *J. Chem. Phys.* **22**, 2045 (1954).
2. P. MOORE, *The Planet Venus*, p. 84. Faber and Faber, London (1957).

and should include the usual abbreviated title of the journal, the volume number, the page number and year in that order.

Footnotes, as distinct from literature references, should be indicated by the following symbols—*, †, ‡, ‡‡, commencing anew on each page; they should *not* be included in the numbered reference system.

8. The language preferred is English, but papers may be published in other languages from time to time. An abstract of not more than 200 words, giving the essential contents of the paper should be provided by the author, if possible in English; but the abstract may be submitted in the language of the paper, for translation into English by the editorial staff.

9. To conserve space, authors are requested to mark the less important parts of the paper (such as records of experimental results) for printing in smaller type. For long papers (over 12 double-spaced typewritten pages) sections which could be deleted without destroying either the sense or the continuity of the paper should be indicated as a guide for the editors.

CONTENTS

NUMBER 1

J. E. SHAW: Geomagnetic and current control of *E*-region absorption 1
R. G. BREENE: The bound free continuum for C^- 10
J. PRESSMAN, F. F. MARMO and L. M. ASCHENBRAND: Artificial electron clouds IV—Thermal ionization study, night time cesium release at 101 km 17
G. DE VAUCOULEURS: Multicolor photometry of Mars in 1958 26
H. E. LAGOW, R. HOROWITZ and J. AINSWORTH: Arctic atmospheric structure to 250 km 33
E. A. KREIKEN: The fall-out of meteoric iron particles 39
J. E. SHAW: Outline of a theory of magnetic separation of auroral particles and the origin of the S_D field . 49
J. E. SHAW: A theory of the origin and geomagnetic control of two types of high latitude E_s 56
L. WALLACE and J. W. CHAMBERLAIN: Excitation of 0_2 atmospheric bands in the aurora 60

RESEARCH NOTES

A. and E. VASSY: Relative intensities of Na lines in the emission of sodium ejected from rockets 71
W. C. HOFFMAN: The current-jet hypothesis of whistler generation 72
J. W. CHAMBERLAIN: On the polarization of the oxygen red line in aurorae 73
J. F. LOVERING: Micrometeorite impacts to an altitude of 135 km 75

NUMBERS 2/3

W. E. GROTH and H. v. WEYSSENHOFF: Photochemical formation of organic compounds from mixtures of simple gases 79
H. N. RUNDLE: Ionization of a static interplanetary gas and expected emission lines from this gas . . . 86
R. V. BHONSLE and K. R. RAMANATHAN: Magnetic storms and cosmic radio noise on 25 Mc/s at Ahmedabad (23°02'N; 72°38'E) 99
A. R. THOMPSON and A. MAXWELL: Solar radio bursts and cosmic rays 104
L. ZHEKULIN: Radio wave propagation characteristics of a simple ionospheric model based on rocket data . 110
M. DUBIN: Meteoritic dust measured from Explorer I 121
J. M. MALVILLE: Narrow hydrogen emission in the aurora 130
G. G. BOWMAN: Further studies of "Spread-*F*" at Brisbane—I. Experimental 133
G. G. BOWMAN: Further studies of "Spread-*F*" at Brisbane—II. Interpretation 150
J. W. CHAMBERLAIN and C. SAGAN: The origin of nitrogen ionization in the upper atmosphere 157
S. F. SINGER: Properties of the upper atmosphere and their relation to the radiation belts of the earth . . 165
F. F. MARMO, J. PRESSMAN, E. R. MANRING and L. ASCHENBRAND: Artificial electron clouds—V. Morning twilight study, release of cesium and sodium at 128 and 166 km 174
I. ESCOBAR, V. N. NERURKAR and R. WEIL: Sidereal anisotropy of high cosmic rays—II 187

BOOKS RECEIVED 193

PAPERS TO BE PUBLISHED IN FUTURE ISSUES 194

NUMBER 4

G. G. BOWMAN: Some aspects of sporadic-E at mid-latitudes 195
A. LUNDBAK: Some simple formulas for latitude effects and lifetimes of satellites 212
G. G. BOWMAN: Triple splitting with the *F*2-region of the ionosphere at high and mid-latitudes . . 214
L. D. DE FEITER, A. D. FOKKER, H. P. TH. VAN LOHUIZEN and J. ROOSEN: Solar radio events and geomagnetic storms 223
J. PRESSMAN, F. F. MARMO and L. M. ASCHENBRAND: Artificial electron clouds—VI. Low altitude study, release of cesium at 69, 82 and 91 km 228
J. CROUCHLEY and N. M. BRICE: A study of "chorus" observed at Australian stations 238
A. OMHOLT: The time delay of the red [01] lines in the aurora 246
G. J. F. MACDONALD: Stress history of the moon 249
S. MIYAMOTO: A geological interpretation of the lunar surface 256

RESEARCH NOTE
H. NORINDER: Some comments on the penetration of whistlers through ionospheric layers 261

ERRATUM
S. F. SINGER: Properties of the upper atmosphere and their relation to the radiation belts of the earth . . 263

VOL 2 1959

AUTHOR INDEX TO VOLUME 2

AINSWORTH, J. 33
ASCHENBRAND, L. M. 17, 174, 228

BHONSLE, R. V. 99
BOWMAN, G. G. 133, 150, 195, 214
BREENE, R. G. 10
BRICE, N. M. 238

CHAMBERLAIN, J. W. 60, 73 (R), 157
CROUCHLEY, J. 238

DUBIN, M. 121

ESCOBAR, I. 187

FEITER, L. D. DE. 223
FOKKER, A. D. 223

GROTH, W. E. 79

HOFFMAN, W. C. 72 (R)
HOROWITZ, R. 33

KREIKEN, E. A. 39

LAGOW, H. E. 33
LOHUIZEN, H.P. TH. VAN. 223
LOVERING, J. F. 75 (R)
LUNDBAK, A. 212

MACDONALD, G. J. F. 249
MALVILLE, J. M. 130
MANRING, E. R. 174
MARMO, F. F. 17, 174, 228
MAXWELL, A. 104
MIYAMOTO, S. 256

NERURKAR, N. 187
NORINDER, H. 261 (R)

OMHOLT, A. 246

PRESSMAN, J. 17, 174, 228

RAMANATHAN, K. R. 99
ROOSEN, J. 223
RUNDLE, H. N. 86

SAGAN, C. 157
SHAW, J. E. 1, 49, 56
SINGER, S. F. 165

THOMPSON, A. R. 104

VASSY, A. 71 (R)
VASSY, E. 71 (R)
VAUCOULEURS, G. DE. 26

WALLACE, L. 60
WEIL, R. 187
WEYSSENHOFF, H. V. 79

ZHEKULIN, L. 110

VO
2
195

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 1–9. Printed in Great Britain.

GEOMAGNETIC AND CURRENT CONTROL OF *E*-REGION ABSORPTION

J. E. SHAW

Ionospheric Prediction Service, Department of the Interior, Miller's Point, Sydney, N.S.W.

(*Received* 20 *June* 1959)

Abstract—Seasonal variations of ionospheric absorption are studied from the mid-day routine tabulations of the parameter f_{min}, from eighteen ionospheric stations. A quantity, $B = f_{min} : f_0E$, is defined and its variations in both geographic and geomagnetic latitudes are studied. It is shown that the value of B varies about a steady value, the amplitude and sense of the variation being related to the intensity of the current in the S_q current system and the latitude of the station. A lunar semi-diurnal variation in the quantity B is found for Kihei. The ratio of the solar and lunar variations in B is found to compare favourably with the ratio of the total currents in the S_q and lunar current systems.

A recent model of the *E*-layer is used to define the upper and lower limits of the absorbing region, and it is concluded that the absorbing region occupies the whole of the dynamo-current region below the reflection height of the exploring waves.

The control of the absorption process is related to the vertical drift of neutral ionisation resulting from east–west current flow. A mechanism is proposed in which the scale size of scattering irregularities in the *E*-layer is changed by the direction and velocity of the vertical drift of ionisation.

A number of observed absorption effects are discussed and it is shown that the theory describes the qualitative behaviour observed in these effects. Reference is made to the likely importance of the process of control of absorption in the auroral zones.

1. INTRODUCTION

Measurements of ionospheric absorption of short radio waves have been made at a number of places on the earth during the past 20 years. Of these those made at Slough since 1935 have been analysed by Appleton and Piggott[1]. These authors derive the relationship

$$-\log \rho = A\,(f+f_L)^{-2} \qquad (1)$$

to describe the absorption, where ρ is the effective ionospheric reflection coefficient, f_L is the longitudinal gyro-magnetic frequency, A is a constant at the time of observation and f is the observing frequency (not too near the critical frequency of any region under observation).

The observations showed that the relation did not hold in winter and it was suggested that the anomaly was connected with the behaviour of the *D*-region at low values of solar zenith angle.

More recently, it has been shown by Whitehead[2] that the absorption values obtained at Slough on frequencies of 2 Mc/s and 4 Mc/s can be better described by the relationship

$$-\log \rho = C + B\,(f+f_L)^{-2} \qquad (2)$$

where

(a) C varies through the seasons with solar zenith angle χ.

(b) C does not alter appreciably with the epoch of the solar cycle.

(c) B seems to depend mainly on the season and not in any obvious way on χ.

(d) Winter variations in absorption are described largely by variations in B.

In this paper it is shown that world wide variations in absorption, apart from those directly dependent on the solar zenith angle χ, are controlled by the current density of the quiet solar and lunar semi-diurnal current sys-

tems. In particular, for absorption measurements made by means of reflections from the E-region, the variations are proportional to the current intensity below the reflection level. Results obtained during the analysis, together with experimental observations made in many parts of the world, suggest that absorption is further controlled by the strength of the S_D and D_{st} currents in the ionosphere. The physical effect is interpreted in terms of the control of scattering phenomena in the lower E-region by means of the vertical drift of neutral ionisation. Some deductions are made regarding the importance of this effect in auroral regions.

2. THE EXPERIMENTAL DATA

The parameter $f_{\min}$ has been used throughout the analysis. Since July 1957, this parameter has been regularly tabulated at a number of ionspheric stations. $f_{\min}$ is defined as the minimum frequency at which an echo is received from any of the ionospheric layers under observation for pulses at vertical incidence. Because the analysis has been confined to values of $f_{\min}$ determined at the hours 1100, 1200, 1300, local time, it can be assumed that the absorption measured is that of pulses reflected at the E-layer. Occasional solar flare occurrences produce values of $f_{\min}$ greater than the critical frequency of the E-layer and where this has occurred, the value has been reduced to a value below the E-layer critical frequency.

The measure of absorption so obtained depends on a number of factors. These are:

(a) Those peculiar to the equipment.

(b) The intensity and type of man-made interference in the locality.

(c) The intensity and frequency range of noise received from atmospherics.

(d) The energy loss of the wave in travelling through the D- and E-regions.

(e) The energy loss associated with reflection conditions in the E-layer.

Reliable intercomparison of measurements at one station or at separate stations is made difficult by the operation of the factors (a), (b) and (c) listed above. However the analysis has been designed to minimise these effects and shows that there is sufficient consistency in the mean behaviour to provide valuable evidence on the physical processes involved. The effect of each of these factors on the analysis will be discussed briefly.

(a) At the eighteen observatories whose published results have been used, a number of ionosondes of varied design and power output are operating. In spite of this, the conditions of overall sensitivity which provide the best routine ionograms have been standardised within certain limits; a fact which may be expected to provide some consistency in the observed changes in absorption at each station.

(b) Transmission on the medium wave broadcast band at frequencies below 1·6 Mc/s, provides the most serious source of interference at most sites. $f_{\min}$ for the midday hours is normally well above this frequency except at the higher latitude stations, where fortunately this source of interference has been largely avoided.

(c) Atmospherics will most seriously affect measurements at some tropical sites where the equipment uses receivers with long time constant automatic gain control. For the purposes of this analysis the effect is not serious.

(d) Non-deviative absorption of waves passing through a Chapman-like D-region will be dependent on the solar zenith angle χ. An expression of the form $A = C(\cos\chi)^n$ has been used to describe the variations of the E-layer maximum ionisation density as well as the variations of D-layer maximum ionisation density. Various values of the exponent n have been determined and it has been assumed in this analysis that n has approximately the same value for the D- and E-regions. The variations of the non-deviative absorption in the D-region have been normalised by using the ratio

$$B = f_{\min} : f_o E. \qquad (3)$$

A study of the variations in B will then be largely independent of the seasonal variations of the non-deviative absorption in the D-region. These variations are included by Whitehead in the quantity C of equation (2).

Table 1. Monthly Mean Values of *B* (1957 - 1958).

Station	Geographic Latitude	Geomagnetic Latitude	Mean Geographic and Geomagnetic	WINTER		EQUINOX		SUMMER
				Dec. Jun.	Jan. Jul.	Sept.	Mar.	Jun. Dec.
Baguio	16·4	5·0	10·7	0·67	0·69	0·82	0·80	0·77
Panama	9·4	20·7	15·7	0·65	0·66	0·79	0·71	0·63
Okinawa	26·3	15·3	20·8	0·76		0·84	0·84	0·81
Kihei	20·8	20·9	20·9	0·64	0·63	0·78	0·69	0·80
San Juan	18·4	30·0	24·2	0·72	0·75	0·72	0·77	
Johannesburg	26·2	27 0	26·6	0·65	0·64	0·69	0·71	0·72
Tokyo	35·7	25·5	30·6	0·74	0·64	0·90	0·62	0·76
Brisbane	27·5	35·7	31·6	0·70	0·65	0·84		
Capetown	34·2	32·9	33·6	0·61		0·64	0·56	0·63
Watheroo	30·3	41·7	36·0	0·51	0·56	0·70	0·60	0·55
White Sands	32·3	41·4	36·9	0·57	0·59	0·71	0·57	0·65
Canberra	35·3	44·0	39·7	0·49	0·51	0·58	0·51	0·69
San Francisco	37·5	43·6	40·6	0·59	0·58	0·69	0·59	0·65
Washington	38·7	50·1	44·4	0·71	0·63	0·68	0·63	0·69
Christchurch	43·5	48·0	45·8	0·68	0·69	0·59	0·62	0·64
Slough	51·5	54·3	52·9	0·55	0·53	0·43	0·48	0·35
Moscow	55·5	51·2	53·4	0·54	0·61	0·51	0·42	0·31
Kjeller	60·0	60·0	60·0	0·74	0·77	0·54	0·62	0·36

(e) Deviative absorption at reflection in the *E*-region may be considered negligible for most of the f_{min} values used, because they are well below the *E*-layer critical frequency. Where high values of f_{min} occur, they have been limited to 0·9 (f_0E) in the analysis.

3. WINTER VARIATIONS IN *B*

Winter values of the monthly mean *B* have been plotted against the three sets of latitude coordinates listed in Table 1 and the result is shown in Fig. 1(a), 1(b) and 1(c). An inspection of Fig. 1 shows that the *B*-values appear

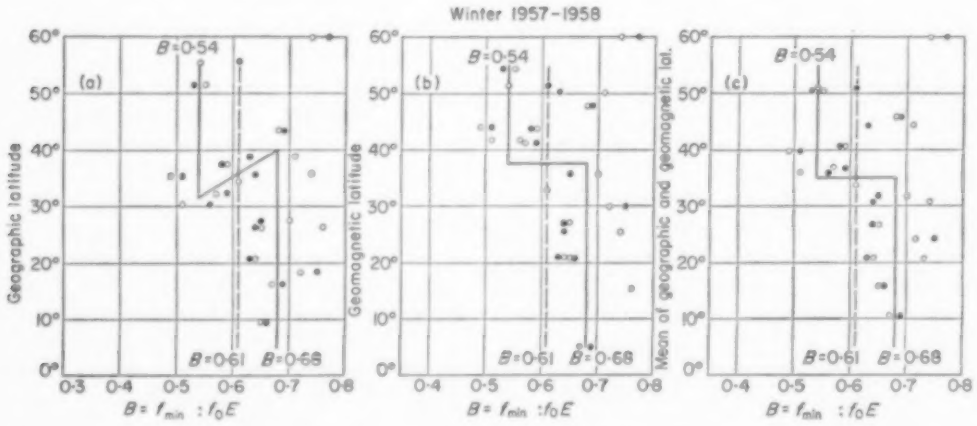


Fig. 1. Winter distribution of monthly mean *B* in (a) geographic latitude. (b) geomagnetic latitude. (c) mean of geographic and geomagnetic latitudes. Legend: ○ December 1957, June 1958. ● January 1958, July 1958.

in two main groups, whose centres have been marked approximately by the heavy lines $B=0{\cdot}68$ and $B=0{\cdot}54$. The overlap in the groups around geographic latitude 35 degrees is removed when the points are presented in geomagnetic latitude (Fig. 1b), indicating that there is some degree of geomagnetic control of B. This separation at latitude 35 degrees is maintained when the mean of the geographic and geomagnetic latitudes is used.

If the discontinuity near latitude 35 degrees is identified with the focal point of the S_q current system in the ionosphere, the behaviour of B on either side of this latitude is just what would be expected if B depended on the direction of current flow in the S_q system. Now the B-values used have been obtained from the midday hours and following Chapman and Bartels[3], it is noted that at this time of day the current in latitudes greater than 40 degrees has a major westerly component, while the current in latitudes less than 30 degrees has a major easterly component. It has therefore been assumed that B is reduced by the presence of a westerly component of the S_q current flowing transversely to the horizontal component of the earth's magnetic field. For an easterly component of current flow, B is increased. Using this assumption the value $B=0{\cdot}61$ has been taken to represent the "no current" value of B and has been indicated in the figures by a dashed line.

4. EQUINOCTIAL AND SUMMER VARIATIONS IN B

If now the equinoctial and summer variations are examined in Figs. 2 and 3, they are shown to be consistent with the assumption made above. At geomagnetic latitudes less than 30 degrees, B is increased in the sense which would be expected from the seasonal increase of current intensity in the S_q system, while at geomagnetic latitudes greater than 50 degrees, B is decreased. Between latitudes 30 degrees and 50 degrees there is a group of B-values which cluster about the "no current" line $B=0{\cdot}61$. This grouping is better defined in Figs. 2(b) and 3(b), and is consistent with the equinoctial and summer variability in the latitude of the S_q current focus as shown by Hasegawa[4]. The mean values of the groups of B-values for equinox have been indicated approximately by heavy lines $B=0{\cdot}75$ and $B=0{\cdot}46$. For summer, the corresponding values are $B=0{\cdot}78$ and $B=0{\cdot}34$.

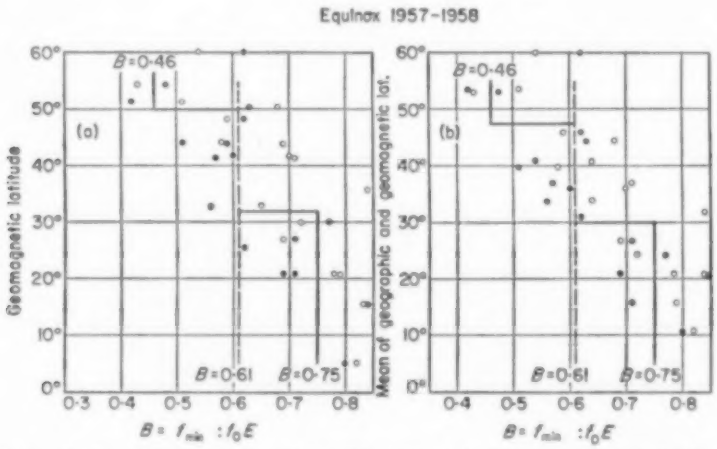


Fig. 2. Equinoctial distribution of monthly mean B in (a) geomagnetic latitude. (b) mean of geo- graphic and geomagnetic latitudes. Legend: ○ September 1957. ● March 1958.

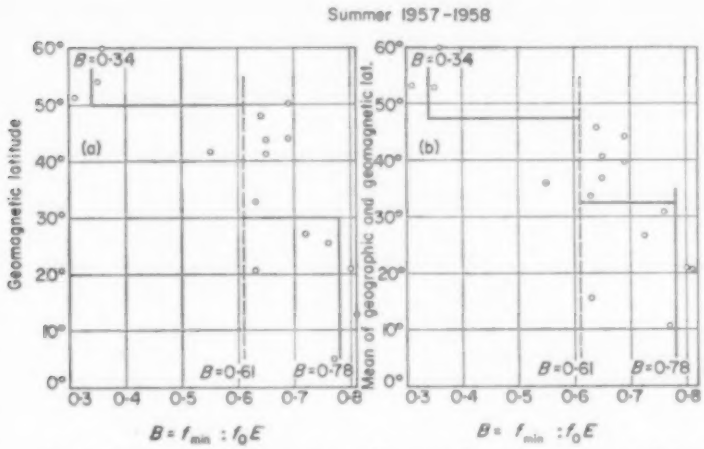


Fig. 3. Summer distribution of monthly mean B in (a) geomagnetic latitude. (b) mean of geographic and geomagnetic latitudes. Legend : O December 1957. O June 1958.

5. LUNAR VARIATIONS IN B

The magnitude of the variations in B which have been assumed to be caused by the S_q dynamo system, suggests that it should be possible to find a lunar variation in B using the results of observations made over a long period at one station. Appleton and Beynon[5] have reported a lunar semi-diurnal tide in absorption measurements made at Slough (geomagnetic latitude 54·3 degrees). A station in low latitudes (Kihei, geomagnetic latitude 20·9 degrees) was chosen so that the phase of the lunar tide, if found, could be compared with that at Slough.

Hourly values of f_{min} for 1200 hr and 1300 hr, taken from 26 months of observations from January 1956 to June 1958, were used. B was determined for each hourly value using the monthly median value of f_0E. The following analysis was then carried out separately for values obtained at 1200 and 1300 hr.

Daily values of B were arranged in columns of 29 and 30 alternately, to give a mean value of 29·5 days or one lunation. Thirty lunations were completed with the data in this way. The 10 international quiet days for each month were then chosen and the mean B for each row in the array was obtained for quiet days and the remaining days. The 29 quiet-day means were then divided into two groups of fourteen and fifteen values, each group covering approximately one cycle of a semi-diurnal oscillation. These two columns were placed next to each other and next to two columns obtained from the quiet-day values at 1300 hr in the same way. The phase of the 1300 hr values was retarded by one day relative to the 1200 hr values to compensate for the phase difference between the hours of observation. There resulted four columns of values representing $(2\pi \times 15)/14{\cdot}75$ radians of any semi-diurnal tidal oscillation present. The fifteen means obtained from these columns are plotted in Fig. 4(a). The analysis was repeated for the remaining days and the results are plotted in Fig. 4(b).

A semi-diurnal oscillation is apparent in the analysis for quiet days. The sinoidal curve which has been fitted approximately, shows an amplitude in B of 0·025 with maximum absorption at about 0300 lunar hr and minimum absorption at 0900 lunar hr. The phase of this variation differs from that at Slough by approximately 8 lunar hr and provides strong evidence that the lunar semi-diurnal variations in absorption are dependent on the current intensity in the lunar current system.

The results of the analysis for all days minus quiet days are confused and this would be

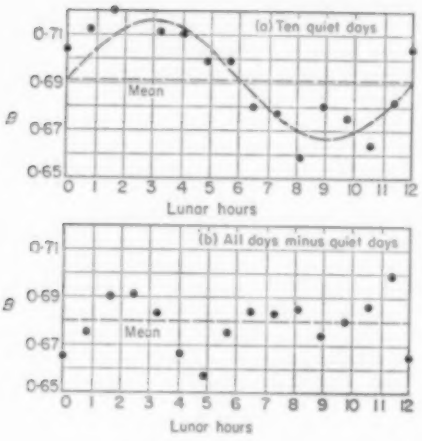


Fig. 4. Semi-diurnal lunar oscillation in B at Kihei, from 30 lunations. (a) Quiet days. (b) All days minus quiet days.

expected if the distribution of disturbance over the period of the analysis is uneven. There is a significant difference in the final means of the two sets of data: $B=0{\cdot}691$ for the quiet days and $B=0{\cdot}68$ for the remaining days. It is pointed out that this difference is consistent with the depression in B which should result from the westward flowing D_{st} currents over Kihei.

The amplitude of the lunar variation in B may be compared with the amplitude of the S_q variation in B at Kihei. It will be assumed that the lunar variation found is principally due to the summer currents and a comparison will be made with the summer variation of B from the "no current" value of 0·61. The ratio of the solar : lunar variation is 7·5 : 1. Now Chapman and Bartels[3] give a total of 11,000 A in the summer lunar current circuit, compared with 89,000 A in the summer S_q current circuit. The corresponding ratio is 8·1 : 1. The two ratios calculated are in fair agreement and if any difference in the height distribution of the solar and lunar currents exists, this may indicate that a larger proportion of the solar current flows below the reflection height in the E-layer. The results of a more refined series of absorption and magnetic measurements should provide valuable evidence on this point.

6. THE AMPLITUDE OF THE S_q VARIATIONS IN B

It will be assumed that the amplitude of the variations in B from the "no-current" value of 0·61 is related to the east-west S_q intensity as follows

$$\Delta B = D \cdot j_x \qquad (4)$$

where D is a constant and j_x is the positive eastwards current intensity of S_q.

Insufficient magnetic data are available to the author to examine (4) fully, but the figures for the seasonal magnetic variations at two representative magnetic observatories will be used. These are Pavlovsk (magnetic latitude 56°) and Bombay (geomagnetic latitude 9·8°) and the figures quoted for the variations in the north component of the magnetic field in units of γ, are those obtained by Schmidt and quoted in reference 3.

Table 2. Seasonal Variations in B and S_q.

Quantity	Winter		Equinox		Summer	
	Value	Ratio	Value	Ratio	Value	Ratio
*Pavlovsk ($C1+C2$)	7·5	1·0	21·0	2·8	25·6	3·4
ΔB (55°)	0·07	1·0	0·15	2·1	0·27	3·9
*Bombay ($C1+C2$)	24·1	3·2	30·3	4·0	30·7	4·1
ΔB (20°)	0·07	1·0	0·14	2·0	0·17	2·4

* $C1$ and $C2$ are the amplitudes of the diurnal and semi-diurnal components of variation in the northward magnetic field.

Examination of Table 2 shows that the relative seasonal changes in B are consistent with the relative seasonal changes in east–west current intensity in S_q at high and low latitudes. It

is apparent, however, that the magnitude of the change in B for a given current flow is much less at latitude 20 degrees than at latitude 55 degrees. It must be assumed then, that a greater proportion of the total S_q current flows below the E-region reflection level at latitude 55 degrees.

Referring to the E-region model calculated by Houston[6] it is found that this behaviour is to be expected. In this model, the height of reflection of a wave whose frequency is well below the critical frequency of the region varies only slightly for values of solar zenith angle $\chi<41$ degrees. For $\chi>41$ degrees the height reflection increases rapidly with increase in χ and it follows that in all seasons, a greater proportion of the S_q current intensity will flow below the E-region reflection height in high latitudes. Equation (4) therefore becomes

$$\Delta B = D\alpha(\chi, h)j_z \qquad (5)$$

where $\alpha(\chi, h)$ gives the proportion of the total current intensity flowing below the reflection height in the E-region and is a function of solar zenith angle χ and the height distribution of the S_q current system.

Kane[7] has reported the results of rocket measurements at Fort Churchill, geographic latitude 58 degrees north. These measurements show that the electron collision frequencies in the 95 kilometer region are less by approximately one third than the values due to Nicolet[8]. The effective conductivity below the E-region reflection heights may therefore be greater in high latitudes than in low latitudes. This would further increase the proportion of S_q current flowing below E-region reflection heights and so modify the value of $\alpha(\chi, h)$ at any given latitude.

7. THE MECHANISM OF ABSORPTION VARIATION

To describe the variations in B, a mechanism is required which depends on the intensity of eastward and westward currents in the region of absorption. Baker and Martyn[9] have shown that the vertical drift of neutral ionisation is controlled by the horizontal current density flowing transverse to the magnetic field in any region of the ionosphere. It is proposed that the process must be explained in terms of this vertical motion of ionisation.

From the phase of the B-variations at midday, it is evident that upward drift of ionisation results in an increase in absorption, while downward drift results in a decrease of absorption. Broad limits are set on the region in which the absorption takes place by the fact that Baker and Martyn have shown that the vertical mobility of ionisation becomes rapidly small below heights of 100 km except in the immediate vicinity of the magnetic equator. The effective conductivity of the ionosphere for the dynamo currents also becomes small, below this height, because of the increase in electron collision frequency. Further evidence on the position of the lower limit of the region is obtained from Houston's model of the E-region. Consider the summer value of B at a latitude of 20°. If $\chi=0$ and $B=0{\cdot}78$, then the height of reflection from the E-region is 109 km. Reference to Table 2 shows that the equinoctial–summer variation of B is small at this latitude and also that $\alpha(\chi, h)$ is small below a height of 109 km. It is concluded that the absorbing region begins at a height of approximately 100 km and occupies the whole region in which the S_q current system is flowing.

Appleton and Lyon[10] and Kirkpatrick[11] have investigated the effects of vertical drift on the distribution of E-layer ionisation. In view of the work of these authors, it is unlikely that the changes in non-deviative absorption resulting from distortion of the E-layer by vertical drift of ionisation will be sufficient to explain the observed changes in B.

It is therefore proposed that the mechanism involves changes in the scattering behaviour of regions of non-uniform ionisation, accompanied by the normal non-deviative absorption processes. Booker[12, 13] has discussed the presence of turbulence in the ionosphere and its importance in a number of scattering phenomena. His theory predicts that there exists a critical value of the scale size of scattering centres, $\lambda/4\pi$, where λ is the wavelength of the scattered radiation. This scale size represents the transition from approximately omni-directional scat-

tering to a predominance of forward scattering. The relatively sharp change from absorption to reflection at the measured f_{min} frequency is now interpreted as a change of the scale size of the scattering centres through the critical value $\lambda/4\pi$, as the frequency of the ionosonde changes through the f_{min} value. For frequencies where omni-directional scattering predominates, the absorption of a wave before and after reflection will be high. When the wave encounters mainly forward scattering on entry into the E-region and on emergence, the absorption will be much reduced.

The variations in B show that the absorption is controlled by the current intensity below the level of reflection. The scale of the irregularities must therefore be changed by the vertical drift of ionisation. Martyn[14] has discussed such a process to explain the behaviour of radio star scintillations and sporadic E-ionisation. Following Martyn, when the ionisation is drifting upwards, any inhomogeneities in ionisation present will be maintained, the surface density of ionisation will be increased and the irregularity will move against the upward drift in a direction determined by the magnetic field, provided that the rate of accumulation of ionisation due to drift is greater than the rate of loss of electrons by recombination. The motion is hydrodynamically unstable[9] and the irregularities of enhanced ionisation break up into smaller regions. It is this last process of breaking up which will control the scale size of the scattering centres and hence determine the absorption behaviour of the region. The downward movement of the iregularities, though very slow in the 105 km region, maintains the scattering medium below the level of reflection in the E-layer.

When the ionisation drift is downwards, the irregularities will move upwards through the reflection level because:

(a) There is a rapid increase in the vertical ion mobility with height,

(b) the S_q current density increases rapidly with height and

(c) there is no adequate source of ionisation at the lower border of the E-layer.

This process will effectively reduce the density of the scattering centres in the region below the reflection level and hence reduce the absorption of the reflected wave.

8. DISCUSSION

The theory put forward in this paper gives a qualitative explanation of a number of features of ionospheric absorption which have not previously been explained. These will be discussed briefly.

(a) Beynon and Davies[15] and others have remarked on the winter days of high absorption in latitudes greater than 45 degrees. It is evident from Fig. 1 that high noon values of absorption will occur when there is no westward current in the S_q system. The diurnal variation of absorption will show its greatest values when there are eastward currents flowing, whether their source be the S_q current system or the S_D current system. For these latitudes $\alpha(\chi, h)$ will have its greatest values in winter.

(b) Davies and Hagg[16] have reported on the conflicting results obtained when absorption is correlated with K index at noon and at midnight. At noon it is to be expected at Prince Rupert, that the S_D current system in the E-region will be opposed in varying degrees by the S_q current system. The resulting absorption changes will therefore show poor correlation with the S_D variations except when S_D is much greater than S_q. It is significant to the mechanism of the absorption process that these authors have remarked on the association of sporadic E with enhanced absorption.

(c) The occurrence, in association with visual auroral displays, of regions of high absorption together with reflecting regions can be explained in terms of the distribution of currents in the display.

(d) The behaviour of E-region absorption during solar flare occurrences should vary with latitude and with time of day.

(e) The theory explains some features of the winter absorption curves obtained by Mitra and Shain[17] during observations of 18·3 Mc/s cosmic radio noise at Hornsby, Australia.

(f) A combination of absorption and magnetic measurements should give more accurate

quantitative information on the distribution of the S_q and lunar current systems in the lower *E*-region.

Acknowledgments—The author is indebted to Dr. D. F. Martyn for helpful discussion on the physical processes involved.

REFERENCES

1. E. V. Appleton and W. R. Piggott, *J. Atmos. Terr. Phys.* **5**, 141 (1954).
2. J. D. Whitehead, *J. Atmos. Terr. Phys.* **10**, 12 (1957).
3. S. Chapman and J. Bartels, *Geomagnetism*, Vol. I, p. 229. Oxford University Press (1951).
4. M. Hasegawa, *Proc. Imp. Acad. Tokyo* **12**, 88 (1936).
5. E. V. Appleton and W. J. G. Beynon, *Nature, Lond.* **164**, 308 (1949).
6. R. E. Houston, Jr., *J. Atmos. Terr. Phys.* **12**, 225 (1958).
7. J. A. Kane, *J. Geophys. Res.* **64**, 133 (1959).
8. M. Nicolet, *Penn. State Univ., Ionosph. Res. Lab. Memo* 171 (1957)
9. W. G. Baker and D. F. Martyn, *Phil. Trans.* **A246**, 281 (1953).
10. E. V. Appleton and A. J. Lyon, *Rep. Phys. Soc., Phys. Ionosph.* 20 (1955).
11. C. B. Kirkpatrick, *Aust. J. Sci. Res.* **A1**, 423 (1948).
12. H. G. Booker, *J. Geophys. Res.* **61**, 673 (1956).
13. H. G. Booker, *J. Atmos. Terr. Phys.* **8**, 204 (1956).
14. D. F. Martyn, *Nature, Lond.* **83**, 1382 (1959).
15. W. J. G. Beynon and K. Davies, *Rep. Phys. Soc., Phys. Ionosph.* 40 (1955).
16. K. Davies and E. L. Hagg, *J. Atmos. Terr. Phys.* **6**, 18 (1955).
17. A. P. Mitra and C. A. Shain, *J. Atmos. Terr. Phys.* **4**, 204 (1953).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 10–16. Printed in Great Britain.

THE BOUND FREE CONTINUUM FOR C^- *

R. G. BREENE Jr.†

Aerosciences Laboratory, Missile and Space Vehicles Department, General Electric Company, Philadelphia 24, Pennsylvania

(*Received* 13 *April* 1959)

Abstract—The wave functions for the ground state of carbon and the negative carbon ion are determined using our small variation program. The Coulomb potential for the free electron is then computed using the carbon wave function thus obtained, and the subsequent obtention of the *s*- and *d*-wave portions of the free electron wave function is detailed. The dipole length form of the matrix elements is then determined for the various desired values of the free electron momenta under the assumption that the lower state is the complete negative carbon ion wave function while the upper state is the antisymmetric product made up of the free electron plus the complete neutral carbon atom wave functions. Finally a simple calculation using the matrix element results provides the photo-ionization cross section.

1. INTRODUCTION

In the boundary layers of certain types of re-entering satellites, nose cones, etc., we will assuredly anticipate the presence of atomic carbon and free electrons. The actual amount of carbon need not concern us in what is to follow, since that forms a specific which can be introduced at any time. In addition to this possibility of occurrence, there is also the strong probability of the carbon continuum which may result from such a mixture occurring in certain stars[1]. The resulting continuum arises from the transition from the state negative carbon ion to the state neutral carbon atom plus free electron. We have developed a treatment of bound-free continua,[2] and we here propose to consider the specific case of the carbon continuum according to this theory.

Let us detail briefly the general procedure which we will follow. In order to give a complete quantitative description of the C^- continuum we need (1) the electron affinity and (2) the matrix element for the bound-free electric dipole transition as a function of the free electron velocity. The former we shall obtain from experiment. The latter we shall determine theoretically, it constituting our prime concern; in order to accomplish this we shall require three wave functions: (1) the wave function for the neutral carbon atom, (2) the wave function for the free electron in the field of the neutral carbon atom and (3) the wave function for the negative carbon ion. We have developed methods for the calculation of these three wave functions, and these computations we shall detail in what is to follow. Subsequently we will make use of these wave functions in order to compute the photo-ionization cross section for the negative carbon ion.

2. CARBON AND NEGATIVE CARBON WAVE FUNCTIONS

Let us first sketch the method which will be used for the determination of the wave functions for carbon and the negative carbon ion.[3] Our expressions for the radial orbitals may be written down as:

$$R_{1s} = 2Z_1^{3/2} r \exp[-Z_1 r] \tag{1a}$$

$$R_{2p} = (1/2\sqrt{2})\, Z_2^{3/2} Z_5^{3/2} [Z_4 - Z_2 Z_5 r] \times \times r \exp[-Z_5 r/2] \tag{1b}$$

* Based on work performed under the auspices of the U.S. Air Force Ballistic Missile Division.

† Present address: 1234 S. Arlington Avenue, Reno, Navada.

$$R_{2p} = (1/2\sqrt{6})\,\overline{Z_3}^{5/2} r^2 \exp[-Z_3 r/2] \qquad (1c)$$

where, of course:

$$\Psi_{nlm}(r, \theta, \varphi) = (R_{nl}(r)/r)\,\Theta_{lm}(\theta)\,\Phi_m(\varphi) \quad (1d)$$

In equation (1d) the angular portion of the wave function is supplied by the well-known spherical harmonic.

The required wave function normality is obviously inherent in equations (1a) and (1c). Normalization for the $2s$ function and orthogonality of the $1s$ and $2s$ functions are furnished by the parameters Z_4 and Z_5. In addition these parameters provide a degree of control over the node of the $2s$ function.

The expression for the atomic energy may be written as:

$$E = \sum_{nl} g(nl)\, I_{nl} + \tfrac{1}{2} \sum_{k} a^k(lm, l'm')\, F^k(nl, n'l') - \sum_{k}^{n'l'm'} \sum_{nlm} b^k(lm, l'm')\, G^k(nl, n'l') \qquad (2)$$

where these energies are, of course, the diagonal elements of the Hamiltonian:

$$H = \sum_{i=1}^{z} (-\tfrac{1}{2}\nabla_i^2 + Z/r_i) + \sum_{i>j} 1/r_{ij} \qquad (3)$$

in atomic units with:

$$I_{nl} = -\tfrac{1}{2} \int_0^\infty R_{nl}(r)\,[d^2/dr^2 + 2Z/r - l(l+1)/r^2]\, R_{nl}\, dr \qquad (4a)$$

$$F^k(nl, n'l') = \int_0^\infty \int R_{nl}^2(r)\, R_{n'l'}^2(r')\, (r_<^k / r_>^{k+1})\, dr dr' \qquad (4b)$$

$$G^k(nl, n'l') = \int_0^\infty \int R_{nl}(r)\, R_{n'l'}(r)\, R_{nl}(r')\, R_{n'l'}(r') \times (r_>^k / r_>^{k+1})\, dr dr' \qquad (4c)$$

Equations (4a), (4b) and (4c) are the energies of the bare nuclear field, the Coulomb interaction and the exchange interaction, respectively. The a^k and b^k in equation (2) are the result of integrating the angular portion of the expansion of $1/r_{ij}$ over the angular coordinates. These, of course, have been evaluated once for all by Slater[4]. We have programmed the analytic expression for equation (2) obtainable from equations (1) and (4) on the IBM 704[5]. In this expression for the energy, the parameters Z_1, Z_2 and Z_3 are completely independent. Simultaneous variation of all three parameters will lead to a set of values for these parameters corresponding to a minimum of the energy as given by equation (2). This, of course, will correspond to the best analytic wave functions of the type equations (1). In our case, the orbitals of the form equation (1d) thus obtained are combined in an antisymmetric product representable by a single determinant. Let us first consider the ground configurations for the atom and ion and subsequently determine the corresponding single determinant wave functions.

We first consider neutral carbon. We know that this atom is possessed of six electrons, two of these being equivalent $2p$ electrons. The Slater diagram corresponding to neutral carbon is given as Fig. 1c, page 1298 of reference (4). Such an electron configuration can, of course, result in ^1S, ^1D or ^3P configurations, the ^3P being the ground state. A single determinant wave function corresponding to the pure ground state configuration may be represented as:

$$\Psi_c[2p(\overset{++}{01})] \qquad (5)$$

where we have distinguished only the $2p$ electrons. The numbers within the parentheses are the magnetic quantum numbers, while the superior signs represent the spins of the electrons in question. The wave function will be a sixth order determinant whose orbitals we shall determine in a moment.

In the case of the negative carbon ion we have three equivalent p electrons, and, although we shall not trouble to do so, it is quite simple to develop a Slater diagram for this distribution of electrons from which the ground configuration may be determined. It so happens that the ground configuration here is ^4S, and one of the single determinant wave functions yielding the pure ground configuration may be represented as:

$$\Psi_{c^-}[2p(\overset{+++}{-101})] \qquad (6)$$

where the notation is as in equation (5). We

now need but to carry out the variation calculation for the configurations corresponding to equations (5) and (6).

The energy expression for the ground states of carbon and the carbon ion respectively are given by:

$$E = 2I_{1s} + 2I_{2s} + 2I_{2p} + F^0(1s, 1s) + F^0(2s, 2s) +$$
$$+ F^0(2p, 2p) + 4F^0(1s, 2s) + 4F^0(1s, 2p) +$$
$$+ 4F^0(2s, 2p) - \tfrac{2}{25}F^2(2p, 2p) - 2G^0(1s, 2s) -$$
$$- \tfrac{3}{3}G^1(1s, 2p) - \tfrac{3}{3}G^1(2s, 2p) - \tfrac{3}{25}G^2(2p, 2p) \quad (7a)$$

$$E = 2I_{1s} + 2I_{2s} + 3I_{2p} + F^0(1s, 1s) + F^0(2s, 2s) +$$
$$+ 3F^0(2p, 2p) + 4F^0(1s, 2s) + 6F^0(1s, 2p) +$$
$$+ 6F^0(2s, 2p) - \tfrac{3}{25}F^2(2p, 2p) - 2G^0(1s, 2s) -$$
$$- G^1(1s, 2p) - G^1(2s, 2p) - \tfrac{3}{15}G^2(2p, 2p) \quad (7b)$$

Equations (7) have been individually minimized on the IBM 704 using our small variation program. As a result, we obtained the following radial orbitals for carbon:

$$R_{1s} = 2(5{\cdot}679)^{3/2} r \exp[-5{\cdot}679r] \quad (8a)$$

$$R_{2s} = (1/2\sqrt{2})(4{\cdot}051)^{3/2}({\cdot}95544)^{3/2}[1{\cdot}5071 -$$
$$- (4{\cdot}051)({\cdot}95544)r] r \exp[-2{\cdot}026r] \quad (8b)$$

$$R_{2p} = (1/2\sqrt{6})(3{\cdot}073)^{5/2} r^2 \exp[-1{\cdot}537r] \quad (8c)$$

The radial orbitals for the negative carbon ion are given by:

$$R_{1s} = 2(5{\cdot}685)^{3/2} r \exp[-5{\cdot}685r] \quad (9a)$$

$$R_{2s} = (1/2\sqrt{2})(3{\cdot}860)^{3/2}({\cdot}9493)^{3/2}[1{\cdot}4436 -$$
$$- (3{\cdot}860)({\cdot}9493)r] r \exp[-1{\cdot}930r] \quad (9b)$$

$$R_{2p} = (1/2\sqrt{6})(2{\cdot}623)^{5/2} r^2 \exp[-1{\cdot}312r] \quad (9c)$$

Equations (8) together with equation (1b) will give the orbitals appearing in the single determinant wave function for the ground state of carbon while equations (9) together with equation (1b) will provide the orbitals required for the single determinant wave function for the negative carbon ion. We now turn our attention to a determination of the free electron wave function.

3. THE WAVE FUNCTION FOR THE FREE ELECTRON IN THE PRESENCE OF THE CARBON ATOM[3]

Our general procedure for determining the wave function for the free electron in the presence of the carbon atom is about as follows: [3] We obtain a potential for the electron by utilizing a single determinant wave function for the ^3P ground configuration of the unperturbed neutral carbon atom. This potential is obtained under the assumption of a Coulomb interaction. In this potential we neglect the effects of (1) exchange forces and (2) forces due to the polarization of the atom by the free electron. In the case of oxygen we have been able to compare the results of these approximations with experiment[2]. In this case the theoretical and experimental results begin to diverge a few hundred angstroms away from the threshold toward the blue. As we proceed toward the blue the experimental results increase slowly while the theoretical results decrease. The experimental results are given down to a minimum wavelength of about four thousand angstroms. At this wavelength the experimental results exceed the theoretical by a factor of very nearly three. Our approximation as to exchange and polarization is surely responsible for at least part of this discrepancy. Even so, the fact that the approximation yields results of this degree of approximation is sufficient to recommend its application to the present situation.

This Coulomb potential we utilize in the Schroedinger equation for the s- and d-waves of the free electron. These s- and d-waves are, of course, the first and third terms in the expansion in terms of Legendre polynomials which we assume as a solution to the free electron Schroedinger equations. The equations are solved numerically and fitted to familiar asymptotic solutions. Only the s- and d-waves will be obtained since this calculation is being carried out with a eye toward specific subsequent bound-free matrix element computations.

Our assumption of a Coulomb potential leads to the following:

$$V_e = -\frac{Z}{r} + \int \frac{|\Psi_e[2p\overset{++}{(01)}]|^2}{r_{er}}\, d\tau \quad (10)$$

In equation (10) the first term is the Coulomb potential due to the carbon nucleus while the second term is the Coulomb potential due to the orbital electrons. The equation appears in atomic units. The wave function indicated is, of course, the six by six determinant for carbon which we obtained in the preceding section. The $r_{c\tau}$ is the separation of the free electron from the volume element $d\tau$ in the orbital charge cloud. We now utilize equations (8b) and (1d) in order to evaluate equation (10). The result is:

$$V_c = -2(5{\cdot}679+\frac{1}{r})e^{-11{\cdot}358r} - 2(6{\cdot}61628r^2 +$$

$$+4{\cdot}64698r+3{\cdot}23258+\frac{1}{r})e^{-6{\cdot}051r} - 2(1{\cdot}20911r^2 +$$

$$+2{\cdot}36079r+2{\cdot}36471+\frac{1}{r})e^{-2{\cdot}072r} \qquad (11)$$

It is certainly apparent that the first term, which is the potential due to the 1*s* electrons may be neglected. Further, there are two terms which are the angularly dependent terms due to the 2*p* electrons, and which have not been included in equation (11). We have shown that these terms are negligible[5], and they have not been included. We now utilize this restricted form of the potential in the fashion of Chandrasekhar and Breen[6] for the hydrogen atom.

We begin by assuming that our free electron wave function is of the form:

$$\Psi_c(r,\theta) = \sum_l \frac{\chi_l(r)}{r} P_l(\cos\theta) \qquad (12)$$

where $P_l(\cos\theta)$ are the Legendre polynomials, and the $\chi_l(r)$ are the unknown radial wave functions for which we must solve the Schroedinger equation. At this point we intend to limit ourselves to a consideration of the *s*- and *d*-wave portions of the free electron solution, since it may be quite readily shown that only these portions of the free electron wave functions combine in the matrix element for the bound free transition. Consequently, we now substitute equation (12) into the Schroedinger equation containing the potential given by equation (11) modified and multiply through on the left by P_0 and P_2 respectively. We subsequently integrate over the polar angle to obtain the following:

$$\frac{d^2\chi_0}{dr^2} + k^2\chi_0 - 2\sum_l \left(\int d\tau_0 P_0 P_l V_c\right)\chi_l \Big/ \int d\tau_0 P_0 P_0 = 0 \qquad (13a)$$

$$\frac{d^2\chi_2}{dr^2} + \left[k^2 - \frac{6}{r^2}\right]\chi_2 - 2\sum_l \left(\int d\tau_0 P_2 P_l V_c\right)$$

$$\chi_0 \Big/ \int d\tau_0 P_2 P_2 = 0 \qquad (13b)$$

It is apparent that all terms in the sum save χ_0 in equation (13a) and χ_2 in equation (13b) drop out.

Numerical solutions to these two equations have been carried out on the IBM 704 in the region $0 \leqslant r \leqslant 8$ for values of the free electron momentum ranging from 0·01 through 0·80 atomic units. The numerical solution for the *s*-wave was joined at the outer end of its range to the familiar asymptotic solution:

$$\chi_0(8) = \frac{1}{k}\sin(k8+\delta_0) \qquad (14)$$

A numerical solution of the *d*-wave remains dependent on the centrifugal potential at the outer end of our range, so it is not asymptotic to an equation of the form equation (14). We know from previous calculations, however, that the phase shift is small so that we may obtain it from the familiar equation:

$$\delta_2 = \int V_c [J_{\frac{5}{2}}(kr)]^2 r dr \qquad (15)$$

Having obtained a phase shift for the *d*-wave we may normalize it by equating it to the following expression at its outer end:

$$\chi_2(r) = \left(\frac{3}{k^3 r^2} - \frac{1}{k}\right)\sin(kr - \pi + \delta_2) -$$

$$-\frac{3}{kr}\cos(kr - \pi + \delta_2) \qquad (16)$$

with which we have effectively obtained the desired free electron wave functions. In Fig. (1) we have plotted the *s*- and *d*-wave functions for the *k* values 0·01, 0·08, 0·10, 0·30 and 0·80 as examples.

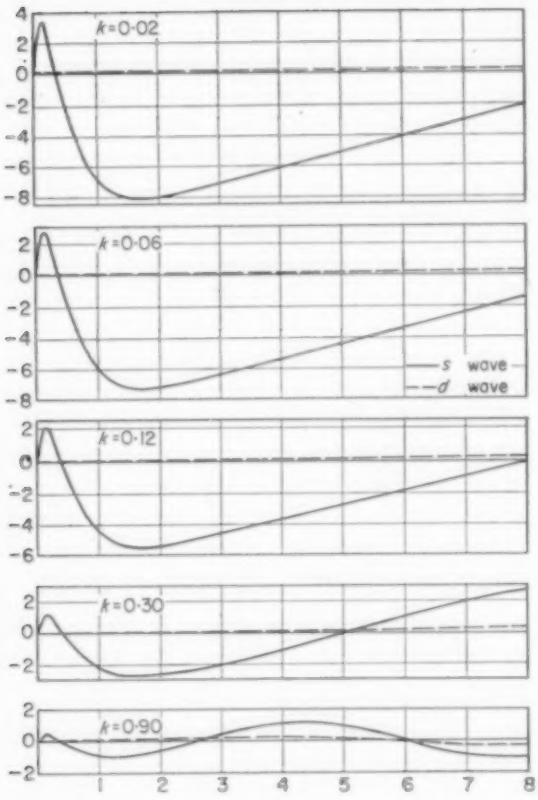


Fig. 1. The s- and d-wave portions of the wave function for the free electron in the presence of the neutral carbon atom for the indicated values of the momentum, k.

4. THE BOUND-FREE CONTINUUM OF C^-

In carrying out this calculation we must determine (1) the form of the initial system neutral carbon plus free electron (2) the form of the final system negative carbon and (3) the form of the electric dipole moment operator which are to be utilized in the calculation.

We have given as our neutral carbon wave function a single determinant of order six or an antisymmetric product made up of six orbital wave functions and containing six factorial terms. What we propose to do is to form an antisymmetric product from these six orbital wave functions and the free electron wave function. We may write this antisymmetric product as:

$$\frac{1}{\sqrt{7!}} P(-) \Psi_{1s\bar{0}}(\mathbf{r}_1) \Psi_{1s\bar{0}}^{+}(\mathbf{r}_2) \Psi_{2s\bar{0}}(\mathbf{r}_3) \Psi_{2s\bar{0}}^{+}$$

$$+(\mathbf{r}_4) \Psi_{2p\bar{0}}^{+}(\mathbf{r}_5) \Psi_{2p\bar{1}}^{+}(\mathbf{r}_6) \Psi_{\epsilon}(\mathbf{r}_7) . \qquad (17)$$

The P in equation (17) is, of course, the permutation operator.

The wave function for the negative carbon ion is likewise given by an antisymmetric product as follows:

$$\frac{1}{\sqrt{7!}} P(-) \Psi_{1s\bar{0}}(\mathbf{r}_1) \Psi_{1s\bar{0}}^{+}(\mathbf{r}_2) \Psi_{2s\bar{0}}(\mathbf{r}_3) \Psi_{2s\bar{0}}^{+}$$

$$+(\mathbf{r}_4) \Psi_{2p\bar{0}}^{+}(\mathbf{r}_5) \Psi_{2p\bar{1}}^{+}(\mathbf{r}_6) \Psi_{2p-\bar{1}}^{+}(\mathbf{r}_7) \qquad (18)$$

We suppose that the incident radiation is unpolarized. Then we would simply average over the matrix elements for the three directions of polarization. Fortunately, a simplification arises here. It turns out that we need compute only the x- and y-polarized radiation; let us detail the reason for this.

An electron is ejected by the incident photon. We consider this electron as immediately establishing a radical electric field, and this field in turn may be taken as establishing a space-fixed z-axis. When we assume this we can show that, in the case of carbon, the matrix elements for the absorption of z-polarized radiation disappear, thus allowing only the absorption of x- and y-polarized radiation. This is not as selective as might be supposed. We may orient our carbon atom in any manner we like before the incidence of the photon, so that the random orientation to be anticipated will not restrict the radiation to be absorbed in the over-all. However, in finding the absorption we will need only to determine the matrix elements for the x- and y-polarized radiation. To further simplify matters the matrix elements of the x- and y-polarized radiation are equal in absolute value. Therefore we shall simply determine the former.

We shall utilize the dipole length form of the

operator which for our case is:

$$M_x = \sum_{j=1}^{7} r_j \sin\theta_j (e^{-i\varphi_j} + e^{-i\varphi_j})/2 \qquad (19)$$

Equations (17) through (19) provide the necessary functions for our matrix element evaluations. Let us first write down the expression for the matrix element and then explain it:

$$(Ce|M_x|C^-) = \sqrt{\frac{2\pi}{3}} \left[\int R_{1s}(C) R_{1s}(C^-)\,dr\right]^2 \left[\int R_{2s}(C) R_{2s}(C^-)\,dr\right]^2 \left[\int R_{2p}(C) R_{2p}(C^-)\,dr\right]^2 \times$$
$$\left[\int R_{2p}(C^-)\chi_0 r\,dr - \frac{1}{5}\int R_{2p}(C^-)\chi_2 r\,dr\right] + \sqrt{\frac{2\pi}{3}} \left[\int R_{1s}(C) R_{1s}(C^-)\,dr\right] \left[\int R_{2s}(C) R_{2s}(C^-)\,dr\right]^2 \times$$
$$\left[\int R_{2p}(C) R_{2p}(C^-)\,dr\right]^2 \left[\int R_{1s}(C^-)\chi_0 dr\right] \left[\int R_{1s}(C) R_{2p}(C^-)\,r dr\right] + \sqrt{\frac{2\pi}{3}} \left[\int R_{1s}(C) R_{1s}(C^-) dr\right]^2 \times$$
$$\left[\int R_{2s}(C) R_{2s}(C^-)\,dr\right] \left[\int R_{2p}(C) R_{2p}(C^-) dr\right]^2 \left[\int R_{2s}(C^-)\chi_0 dr\right] \left[\int R_{2s}(C) R_{2p}(C^-)\,r dr\right] \qquad (20)$$

We first remark the approximations which have been made. We have assumed that the 1*s* orbital for carbon is orthogonal to the 2*s* orbital for negative carbon and that the 2*s* orbital for carbon is orthogonal to the 1*s* orbital for negative carbon. This is very nearly true and should certainly not affect our results. Let us next consider the first term in the three term sum of equation (20). In this combination of wave functions in the integrands involved, we represent the possible case where the free electron proceeds to a 2*p* orbital under the aegis of the dipole moment operator. Since this operator is a one particle operator and we have assumed orthogonality, all other orbitals in the carbon atom combine only with the same orbitals in the carbon ion. Now there are seven factorial terms in the antisymmetric product for the initial system and seven factorial terms in the product for the final system. Each member of one of these products will combine in an integral identical to this first one with a single member of the other product. Thus, there will be seven factorial integrals precisely similar to the first term in equation (20). Division of the sum by the normalization factor yields a coefficient of unity. The second term in equation (20) represents the situation wherein the *s*-wave portion of the free electron wave function combines with the 1*s* orbital while a 1*s* orbital from neutral carbon becomes a 2*p* orbital in the negative carbon ion. The last term is similar to the second except that now we consider our free electron as proceeding to a 2*s* orbital while one of the 2*s* electrons in the carbon atom proceeds to a 2*p* orbital in the negative ion. These represent all the possibilities for combination among the twice seven factorial members of the two antisymmetric products. And now it is simply a matter of numerical integration to combine our tabular free electron wave functions with our analytic bound orbital functions.

We have programmed this calculation for the IBM 704 and carried it out for carbon for the various values of the free electron momentum for which wave functions were determined. We next compute the photo-ionization cross-section.

From Chandrasekhar[7] we obtain the following expression for the cross-section in question:

$$\sigma = 6{\cdot}812 \times 10^{-20} k (k^2 + 2I) |(Ce|M_x|C^-)|^2. \qquad (21)$$

Let us recall that the absorbed radiation will correspond in energy to the free electron energy plus the electron affinity of carbon. This electron affinity we shall take as 1·12 electron volts[1]. We now have the two factors necessary to our continuum, namely, photo-dissociation cross-section as a function of radiant energy. We have plotted this continuum as Fig. 2. It may be remarked that the cross section computed is of the same order of magnitude as that for oxygen[8].

If one were interested in the specific application of our results to boundary layer radiation, one would, of course, first determine

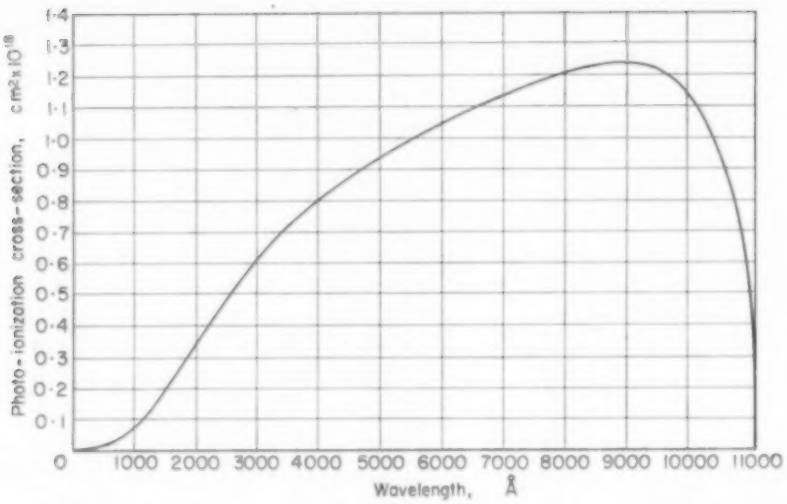


Fig. 2. The cross-section for the photodetachment of an electron from the negative carbon ion.

the absorption coefficient from this cross section. This in turn could then be inserted in a simple emissivity expression of the type: [9]

$$\varepsilon_\nu = 1 - \exp\{[-aL\mu_\nu][1 - \exp(-h\nu/kT)]\}$$

in order to obtain the actual emissivity of a layer of carbon ions of a particular thickness, density and temperature.

Acknowledgements—The author wishes to express his appreciation to Mrs. Rosemary Johnson for carrying out the necessary IBM 704 computations. As always it is a great pleasure to acknowledge the continuing assistance of Miss Maria C. Nardone in this as in many previous calculations.

REFERENCES

1. L. M. Branscomb and B. E. J. Paget, *Mon. Not. R. Astr. Soc.* **118**, 258 (1958).
2. R. G. Breene, Jr., and Maria C. Nardone, *The Bound-Free Continuum of O⁻*, General Electric Company TIS R58SD 296, Philadelphia (1958).
3. R. G. Breene, Jr., *Phys. Rev.* **111**, 1111 (1958).
4. J. C. Slater, *Phys. Rev.* **36**, 51 (1930).
5. R. G. Breene, Jr., and Maria C. Nardone, *Phys. Rev.* **115**, 93 (1959).
6. S. Chandrasekhar and F. H. Breen, *Astrophys. J.* **103**, 41 (1946).
7. S. Chandrasekhar, *Astrophys J.* **102**, 223 (1945).
8. L. M. Branscomb, D. S. Burch, S. J. Smith and S. Geltman, *Phys. Rev.* **111**, 504 (1958).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 17–25. Printed in Great Britain.

ARTIFICIAL ELECTRON CLOUDS–IV

THERMAL IONIZATION STUDY, NIGHT TIME CESIUM RELEASE AT 101 km

J. PRESSMAN,* F. F. MARMO* and L. M. ASCHENBRAND*

Geophysics Research Directorate, Air Force Cambridge Research Center,
L. G. Hanscom Air Force Base, Bedford, Massachusetts

(*Received* 15 *May* 1959)

Abstract—There is reported radio-radar and optical observations of an electron cloud created by the night time release of cesium at 101 km from a Nike-Cajun rocket launched at Holloman Air Force Base on 20 May 1958. An analysis of the electron yield based upon the thermochemistry and the Saha relationship indicates that the electrons observed can be reasonably ascribed to thermal ionization.

1. INTRODUCTION

Another paper of this series[1] reported the generation of an artificial electron cloud by the daytime release of atomic potassium at an altitude of 121 km. Although the primary process for the production of electrons was ascribed to the solar photoionization of the contaminant, the observed radar data suggested that some process other than that of photoionization was responsible for the early, initial (first 100 sec) electron production.

Qualitative arguments indicated that it was not unreasonable to identify this early process with the thermal ionization of potassium, the necessary heat being supplied by the heats of chemical reaction of the release method. The primary object of the daytime release of potassium was to ascertain the role of the solar photoionization of the contaminant; thus it was of considerable importance to determine to what extent processes other than photoionization may contribute to the observed electron cloud. In addition, since the series of experiments were to utilize three alkali metals, Cs, K and Na, it was felt that the possibilities of thermal ionization must be examined with all three contaminants.

Concerning thermal ionization only, it is not necessary to study each of the alkali metals in detail. Rather, it is sufficient to utilize Cs and to extend the findings to the cases of potassium and sodium. This is illustrated in Fig. 1 which depicts the behavior of K_n with temperature T, the equilibrium constant for the process:

$$X \rightarrow X^+ + e \quad \text{where } K_n = N^+ N_e / N_0 \qquad (1)$$

and N^+, N_e, and N_0 are the number densities of the ions, electrons and neutral atoms, respectively. Fig. 1 clearly illustrates that, other parameters considered equal at a given temperature T, Cs produces considerably more electrons than the other two alkali metals taken into account. Note that Fig. 1 is a plot of the $\log_{10} K_n$ vs. temperature.

It is expected that the chemistry of generation of the alkali metals is similar in each case (Cs, K and Na). In addition, it can be shown that the thermochemical data associated with alkali metal nitrate–aluminium reactions are similar; thus essentially the same amount of cannister heating is expected in each case. Therefore, in order to ascertain a maximal type of extent of thermal electron production it was felt that cesium was the best choice. Also, since the daytime (solar photoionization process

* At present associated with Geophysics Corporation of America, 700 Commonwealth Avenue, Boston, Massachusetts.

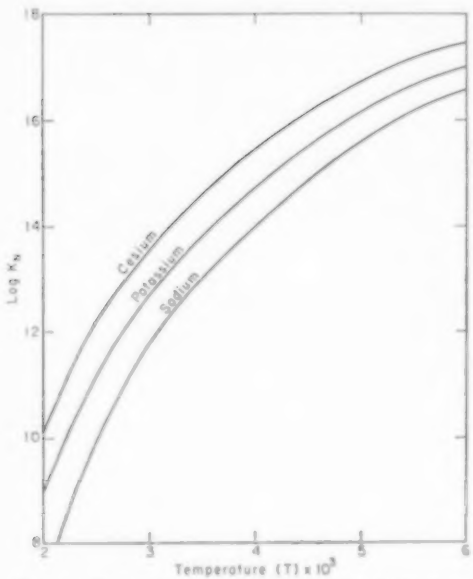


Fig. 1. The variation of the thermal ionization equilibrium constant for the alkali atoms as a function of temperature (°K) calculated from the Saha equation.

of electron production) release of Cs was planned, the above arguments only serve to point up the necessity of determining the extent of thermal production of electrons for the Cs case especially.

The reason for the choice of a nighttime release is of course to isolate, through experimentation, any contribution by solar photoionization and photodetachment processes. However, it should be emphasized that this does not imply that only the process of thermal ionization need be considered for the generation of electrons. Other processes are possible, but here the only consideration is thermal production. The approach will be justified by showing that it is possible and reasonable to ascribe the observed electron production to thermal processes.

The decision to release the Cs vapor at an altitude of 100 km was determined largely by the desirability of producing an easily visible and photographable cloud. The advantages of such a procedure are manifold. No detailed discussion is offered here; instead, only those advantages utilized in the present problem are pointed out: namely, the ascribing of the electron generation to a thermal process. In addition, optical data pertaining only to this end are included. Some of the pertinent advantages gained by optical detection of the cloud are as follows: (1) triangulation optics can yield very accurate (within 1 per cent) altitude data (2) optical determination of the initial cloud size can be correlated with simultaneous radar returns (this can aid in determining the total electron content) (3) change of size with time can associate time dependence of observed radar returns with optical tracking. These illustrate only some of the advantages of the visible cloud technique as considered.

The method which rendered a Cs release photographable was straightforward. It was decided to "dope" the cesium release with about 10 per cent sodium vapor. The sodium vapor would react in a chemiluminescent manner with the atomic oxygen at 100 km (maximum number density of 0 atoms) and produce the sodium *D*-line (5890 Å). The determination of the necessary amount of sodium to utilize (5 per cent) and the choice of altitude (100 km) for maximum emitted intensity were arrived at from reference to the recent rocket release of sodium into the upper atmosphere[2]. Actually, about 13 per cent Na was used as "doping agent" in order to have a working excess to allow for unpredictable parameters such as overall yield, vigor of release and wind shears.

2. EXPERIMENTAL TECHNIQUES

The chemical generation of alkali metals (Cs in this case) and the subsequent release has been described for a similar experiment[1, 3, 4]. One significant difference exists: the choice of vehicle. In the previous experiment an Aerobee rocket was employed. This permitted the use of a cannister of such volume and weight that over 140 moles of atomic potassium could be carried to 121 km. However, in anticipation of a systematic study utilizing many rockets, it was decided to attempt to continue this series using the less expensive, solid fuel Nike-

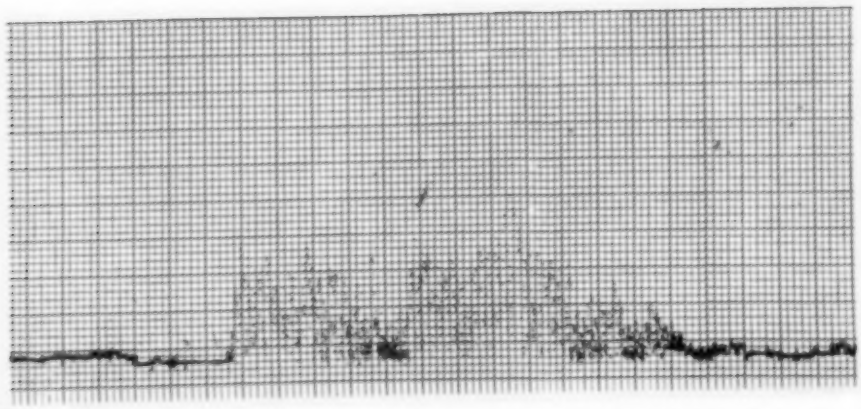


Fig. 2. A typical Dovap record on 36·9 Mc. of 20th May, 1958. The time scale is indicated by the dur-ation of 65 sec. from onset to final decay. The vertical scale is logarithmic running from 0·3 at base to approximately 100.

[facing p. 19

Cajun.[1] The payload–altitude capabilities of this vehicle were considered suitable for this experiment.

A stainless steel cannister to contain the contaminant was designed and constructed to be employed within the requirements of payload, configuration of nose cone, high acceleration (approximately 60 g's), bending movement, and adaptability to the rocket body and instrument package. The vaporizer measured 12 in. in length and 6·5 in. in diameter (volume 400 in^3). The weight of the empty vaporizer was 8·6 kg. The chemicals used for the experiment consisted of 8·4 kg $CsNO_3$ and 4·0 kg of Al granules.

The ratio of gram moles of sodium nitrate (doping agent) to cesium nitrate was 5·4/43·1 or about 13 per cent.

A Nike–Cajun rocket carrying the described loaded cannister was launched at 0243:35 MST, 20 May 1958, at Holloman Air Force Base in Alamagordo, New Mexico. A preset timer actuated the release of atomic cesium and sodium at 92 sec after launch at an altitude of 101 km.

3. EXPERIMENTAL RESULTS: OPTICAL

Several optical probes were utilized in this experiment; only those data required for the content of this paper are presented here. Other data are presented in another report of this series[3]. Brief descriptions of the optical probes are given here.

The Astro Ballistic Camera net consisted of six cameras located at three different sites to afford accurate (within 1 per cent) altitude determination. Two cameras were located at each site and each was equipped with an $f/2·5$ lens. Correlated optical data were obtained by utilizing star background and synchronized (WWV) timing. Communication difficulties hampered operations slightly, but some useful data were obtained. The cameras were programmed to take 5-min time exposures beginning at release time. Good photographs were obtained, and the following data were deduced. The cloud was initially generated at an altitude of 101 km above a location 10·1 km north and 7·6 km west of the launch site. The apparent photographed cloud shape was oblong, and its dimensions were about 1·2 x 2·3 km. It must be noted that this represents a value which is undoubtedly too large since the time exposure of 5 min represents an integrated response in which the relative motion (wind) of the center of gravity plus any diffusive effects are recorded. A reasonable guess for the initial size would be about one half of the smallest dimension or about 0·6 km diameter sphere.

The Baker–Nunn satellite tracking camera was utilized. This instrument is a 3-axis, 20-in aperture, $f/1$ camera and is located at Organ Pass, New Mexico (35 miles from the launch site). It has a motor-operated mechanical drive which provides sequential tracking or series photos on a 250 x 55 mm frame. Furthermore, accurate (WWV) times are displayed on each exposure. A clear photograph was obtained on this instrument for a total of 26 sec, after which the cloud became too diffuse to photograph. The cloud appeared spherical initially and after several seconds began to assume a rhomboidal geometry. Here, the measured apparent initial sphere diameter was only 0·7 km. This was undoubtedly due to the short exposure times involved. Again, this represents an upper limit; however, it is felt that, in the short time factors considered, a value of initial sphere diameter of about 0·6 km is reasonable.

A commercial 35-mm camera was used to obtain an additional measure of initial cloud size. The camera was an Exacta IV with $f/2·8$ lens (50 mm). A time exposure of 4 min beginning approximately at release time (when first visible) resulted in a clear photograph of the cloud which, when reduced to apparent size, was an elongated rhomboid which measured 0·9 x 1·5 km. The image of the cloud was obviously distorted due to wind carry during the 4-min exposure time. However, the agreement with the Astro cameras (5-min exposure) in ratio of axes and absolute sizes is considered good. From these, a value of 0·6 km for the initial cloud size seems reasonable and, thus, is the value adopted in this study. This is considered to be an observed initial sphere diameter.

4. EXPERIMENTAL RESULTS: RADAR

Most of the radar probes utilized in the present study have been described elsewhere[3]. These include (a) a 6-m and 2-m radio amateur net, (b) the Dovap and Doran White Sands missile tracking systems (c) the Laredo–Smythe Associates forward scatter net (d) the Ft.Huachuca–Collins forward scatter net (e) the Stanford University 70-kw pulsed radar (on 23·1 Mc for this experiment) at Alamogordo and (f) the *C*-3 ionosonde located at the White Sands Proving Grounds, New Mexico.

The 6-m net utilized a 50·01 Mc coded transmitter located 300 miles west (Tucson, Arizona) of the launch site. A ring of amateur monitoring stations was located within a 500 km radius of the test area. Amateur participants were requested to tape record their signals and count-down information. A positive signal of the coded transmission was obtained. The coded spectrogram was easily identified with the Tucson transmission and became detectable only after the cesium release at 101 km. A variable beat oscillator was employed to make the observed signals audible. Readable spectrograms were reproduced for a duration of only about 1 min after release, but audible signals are identifiable for about 8 min following release.

The 2-m net was similar to the 6-m net except that transmissions were on 144·01 Mc. Here again, signal enhancement was recorded in El Paso, Texas (about 100 miles from launch site) at 92 sec after launch. The signal was audible for about 4 min. These results obtained on the radio nets are considered to be significant for utilization of artificial electron clouds for propagation schemes. It is felt that this is the first report of audible transmissions beyond line of sight by means of an artificially produced electron cloud.

A representative result of the Dovap and Doran network is depicted in Fig. 2 which is a photograph of the Dovap record on 36·9 Mc of 20 May 1958. The initial onset, duration (65 sec) and subsequent decay are clearly shown.

The Laredo–Smythe forward scatter system consisted of monitoring of the 201 Mc Laredo, Texas, transmissions. Two locations were utilized for forward scatter monitoring. One was situated at 39° 45′ N latitude and 112° 50′ W longitude in north central Utah while the other was at 36° 35′ N latitude and 115° 55′ longitude west of Las Vegas, Nevada. Both locations received cloud-propagated signals at 91·6 sec after launch. The time duration was 7 sec for the Las Vegas installation, and 9 sec for the Utah site. This time differential may be ascribed to the cloud geometry. The optical data were not sufficient to verify this.

The 810 Mc transmission from Ft.Huachuca was monitored by Collins Radio Company at Lubbock, Texas, and Encino, New Mexico. The Encino site reported a 2-sec burst on 810 Mc immediately following gas release. The Lubbock site reported negative results.

The Stanford University 70-kw pulsed radar was operated on 23·1 for this experiment. In addition, a 6 Mc radar located in the same area was utilized simultaneously and provided useful data. At gas release time a strong return was noted on 23·1 Mc and saturated for about 50 sec (Fig. 3). Thereafter a decrease in amplitude occurred until about 90 sec after release at which time it was no longer detected on 23·1 Mc. On the other hand, the behavior of the 6 Mc return is quite different. The buildup began at about 50 sec until it reached a maximum at 125 sec whereupon it decayed rapidly over the following 25 sec. It is noted that the buildup of the 6 Mc coincides with the decay of the 23·1 Mc. The observed behavior of these two frequencies suggests that the time required to achieve a sufficiently large radar cross-section to 6 Mc (over 500 m^2) is about 50 sec. The observed decay in each case is associated with the gradual electron density decrease due to attenuating factors such as diffusion, recombination, etc. It is possible that a method of determining cloud dimensions and electron density distribution within the cloud can be ascertained by utilizing refined experimental techniques based on these observations. Such data would be of importance in the study of propagation physics.

The records from the *C*-3 ionograms are summarized in Fig. 4 where the maximum frequency return is plotted against the time after

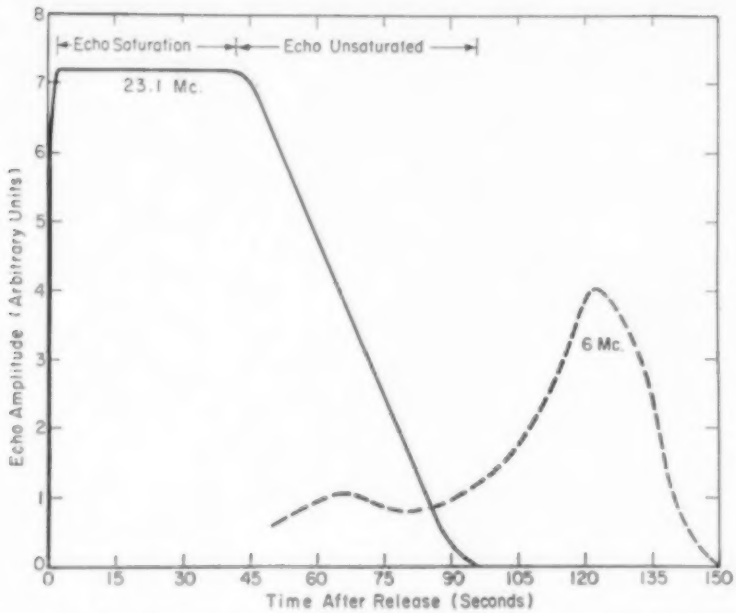


Fig. 3. Echoes from electron cloud by Stanford University 23·1 Mc and 6-Mc radars.

release of contaminant. The sharp onset of high critical frequency noted during the correct 15-sec sweep is evidence of the identification of these and subsequent signals to the newly generated artificial electron cloud. The corresponding electron density is also included for reference. Further evidence is afforded by graphing the apparent slant range with time (Fig. 5). The smooth variation indicates the presence of a small cloud initially detected at the correct time and drifting slowly away from the station. It should be noted that for such a limited time (3 min) no reliable estimates of wind velocity could be attempted.

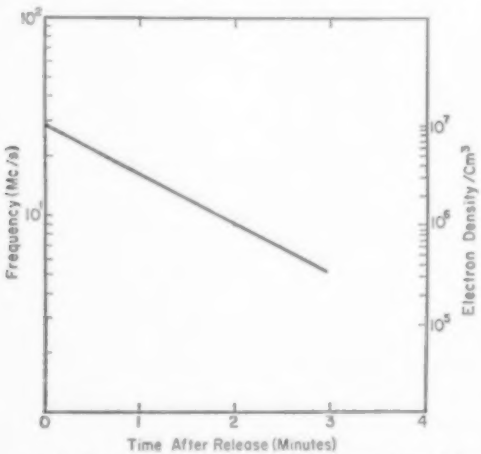


Fig. 4. Plot of maximum frequency return from C-3 as a function of time. The Stanford radar returns data are used during the first 15 secs. for the initial high frequencies and the curve extrapolated to the zero time of release.

It is felt that the generation of a visible, high-density electron cloud was successfully accomplished by the nighttime release of Cs vapor. The fact that the cloud was generated seems indisputable. The phenomena largely responsible for the generation of this cloud has yet to be established with certainty. The remainder of this paper is devoted to this problem. It is shown that it is not unreasonable to ascribe the observed electron cloud to the thermal ionization of the released, high temperature cesium.

5. DISCUSSION OF EXPERIMENTAL RESULTS

The discussion of results is confined to the problem of the electron generation process;

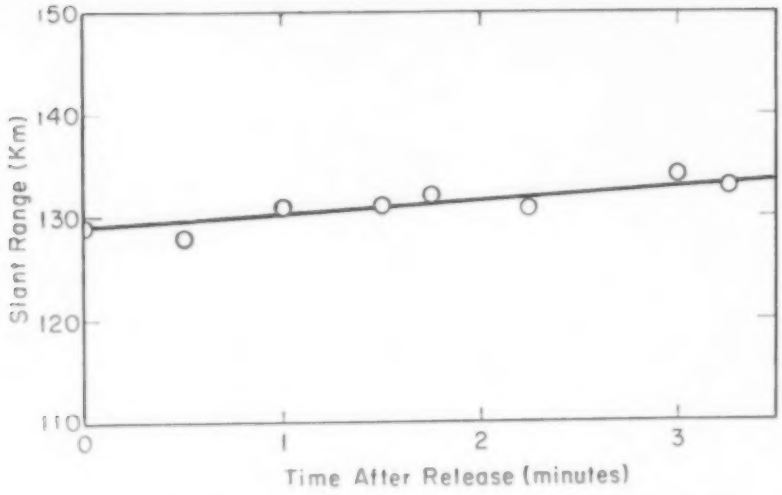


Fig. 5. Variation of slant range of electron cloud with time from C-3 station.

here only the problem of thermal generation is considered. It is clear that for a study of this nature it is essential to be able to estimate the extent of electron generation; that is, to estimate the total electrons generated in the process. Then the simple energetics of the problem can be considered to determine whether enough energy is available to account for the observed result. Finally, if this is satisfactorily accomplished, it then remains to ascribe a model to the process and determine whether or not the generation of the total electrons can be ascribed to a reasonable model by utilizing realistic parameters. This is the approach adopted here. First, an estimate of the number of electrons is required.

The photographic data from the three optical probes (especially the Baker–Nunn camera) indicated that initially the released cloud was a sphere of apparent diameter of about 600 m. Thus, the initial volume is taken to be about 1×10^{14} cm^3. If the following further assumptions are made: (a) the initial density distribution is homogeneous and (b) the yield of Cs production is taken to be about 10 per cent (good to a factor of two), then it is easily shown that the Cs atoms are distributed throughout the sphere with an average density of 3×10^{10} cm^3; this is a factor of 300 less than the density of the total ambient species.

The radar data can be utilized to obtain an estimate of the initial electron density throughout the cloud. Fig. 4 of the *C*-3 data indicates an initial electron density of about 10^7 electrons/cm^3. However, if the Stanford data are extrapolated to zero time, a value closer to about 6×10^7/cm^3 is obtained. Thus an initial electron density value of about 3×10^7 electrons/cm^3 is adopted as realistic for a period of the first few seconds following Cs release. Here, however, the direct correlation between optical and radar observations should be applied with care. For example, while the optical probe may record a cloud of 300 m radius, the "size" of the cloud observed by the radar corresponds to at least that of the radar cross-section which in turn is related to the measured electron density. Thus, it is assumed here that at the radius of reflection the electron density is 3×10^7/cm^3. However, the approximation is again applied that throughout the 300 m (optical) radius of the cloud the local electron density does not vary by a factor of 2 and the average assumed here is that of 3×10^7 cm^3.

Thus if an initial volume of 1×10^{14} cm^3 and initial average electron density of 3×10^7 are

adopted, the total electron content of the artificial electron cloud is estimated to be of the order of 3×10^{21}. Furthermore, if the average density of neutral Cs atoms is initially taken as $3 \times 10^{10}/\text{cm}^3$, then the probability of thermal ionization is of the order of 10^{-3}. It now remains to determine whether the estimates of the parameters discussed above are compatible with ascribing the generation of the electrons to the thermal ionization of the cesium vapor contained in the cannister before release at altitude.

6. THERMAL IONIZATION

Perhaps the most direct mechanism to account for the creation of the observed electron cloud is thermal ionization. Here, it is conceived that the evolved gas is basically processed at a certain temperature and particle density to yield free electrons. It is further supposed that after a plasma is generated, the material is spread outward from the dispenser as it bursts, and the material is diluted in a geometrical fashion in accordance with the observed extent of expansion of the cloud. For example, in this case, the initial sphere of radius is 300 m. Subsequently, the electron concentration decreases in time through attachment, diffusion, recombination and other factors. The theory of thermal ionization is applicable strictly to equilibrium processes. In the present case the governing reaction (discussed below) is very rapid, almost violent in character. Whether an equilibrium or close to equilibrium process is experienced here is somewhat uncertain. Nevertheless, as a first approximation to indicate order of magnitude considerations, the thermal electron yield is calculated for an equilibrium process.

For process (1) we have from the Saha relation[5]

$$\log K_n = \log (g^+ g^- / g^0) - 15{\cdot}38 - 1{\cdot}5 \log T - 5040\, V_i / T \qquad (2)$$

where the g's are the statistical weights, T the absolute temperature and V_i the ionization potential in electron volts. Further,

$$N_e + N_0 = N_0^i \qquad (3)$$

where N_0^i is the total number of original uncharged particles/cm^3; N_e and N_0 have been previously defined. If we assume that $N_e = N^+$, then the following expression is obtained by substitution:

$$2N_e = K_n [(1 + 4 N_0^i / K_n)^{\frac{1}{2}} - 1] \qquad (4)$$

If $N_0^i >> K_n$ (as in the present case):

$$N_e = (K_n N_0)^{\frac{1}{2}} \qquad (5)$$

In order to estimate the total number of electrons produced thermally, we must investigate the thermochemistry of the chemical production of the alkali metal vapor to obtain an estimate of the temperature and the concentration of cesium vapor prior to the cannister burst. The governing reaction is considered to be a mixture of (6) and (7).

$$3\,\text{Al} + \text{CsNO}_3 \rightarrow \text{Cs}\uparrow + \text{Al}_2\text{O}_3 + \text{AlN} \qquad (6)$$

$$2\,\text{Al} + \text{CsNO}_3 \rightarrow \text{Cs}\uparrow + \text{Al}_2\text{O}_3 + \tfrac{1}{2}\text{N}_2 \qquad (7)$$

The energies available for heating the Cs vapor from (6) and (7) are 381 and 224 kcal respectively, allowing 25 kcal for the heat of vaporization of cesium (estimated from sodium data). In addition, it can be shown that a reasonable estimated heat capacity of reaction products is 40 cal/g mole. Because of the speed of the reaction it is assumed that the heat losses due to conduction are negligible. Further, if the reaction is adiabatic, then we obtain, in principal, Cs vapor temperatures of about 9000°K and 5700°K respectively for (6) and (7). Since the reaction is at best considered to be a combination of these reactions, an upper limit of temperature attained can be about 7000°K. However, the actual mechanism is unknown, and the heat capacities and states of reaction products are unknown at the elevated temperatures; therefore any statement concerning the achieved temperature is at best a very crude estimate. It is probably safe to assume that temperatures of the order of 2000°K–4000°K cannot be ruled out.

In order to apply the Saha equation we must obtain a value for N_0^i, the total concentration of Cs atoms. To obtain this value for temperatures between 1000°K–6000°K, the corresponding pressures in atmospheres are calculated taking into account the total number of moles reacted. Here it is assumed that the cesium gas

evolved is at temperature, T, and that it fills the available (void) space in the container at a uniform density. The void space in the container is initially (before reaction) 2·4 l. When (and if) the full 43 moles of $CsNO_3$ is reacted, the available space increases to 4·3 l. For the case of partial consumption of the $CsNO_3$, intermediate volumes are available. If the increasing void space with extent of reaction is allowed for, the pressure, P, for X moles reacted is given by

$$P = \frac{X\,(6 \times 10^{23})\,T}{(2400 + 42X)\,(2{\cdot}7 \times 10^{19})\,273,} \qquad (8)$$

P is in atmospheres, T in °K. The values for X and T for different pressures are plotted in Fig. 6. If one now assumes dynamic bursting pressure (as in the present violent reaction), it is easy to achieve instantaneous pressures greater than the static burst pressure by a factor of 2. After an estimate of the number of neutral Cs atoms/cm³ is obtained, the Saha equation is applied to determine the probability of thermal ionization for the various temperatures for a given burst pressure. Subsequently, the total number of thermal electrons generated and contained in the cannister at burst time is calculated. These data are given in Table I in which the cases of 300 atmosphere and 500 atmosphere bursting pressures are considered, respectively. The ratio of electrons to neutral particles is also given in the table. The temperature range considered was from 1000°K to 5000°K.

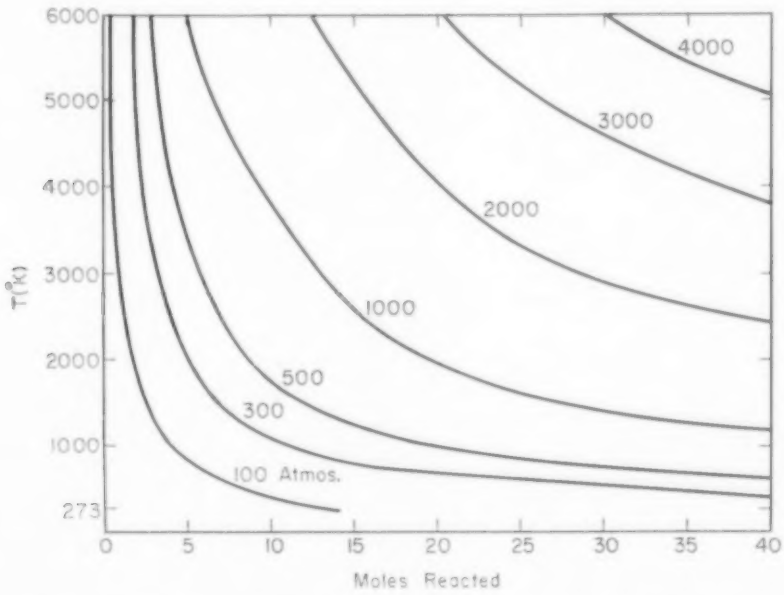


Fig. 6. Pressure within cannister as a function of moles reacted and temperature of the reactants.

pressure of the cannister, the concentration at the burst pressure is found. A burst pressure of about 300–500 atmospheres is adopted here. This value seems reasonable since the static pressure of rupture is estimated to be over 200 atmospheres. For a very rapidly increasing

A variation of burst pressure by a factor of 2 or 3 would not significantly alter the electron concentration for a fixed T since the void space containing these electrons would alter in such a manner as to offset the effect of the pressure factor. In addition, the achieved electron

density varies to the one half power of N_0^i. On the other hand, the temperature is a more sensitive parameter.

With reference to the table, the following values deduced earlier should be recalled here. For example, it was estimated that reasonable requirement values for the total number of electrons and thermal ionization probability were 3×10^{21} electrons and 10^{-3} electrons/neutral atom. Furthermore, a crude estimate of yield was taken as about 10 per cent (within a factor of 2). Table 1 suggests that these values are best satisfied by a temperature above 3000°K and a bursting pressure between 300–500 atmospheres. It is interesting to note that the ratio of electrons to neutrals is only satisfied for temperatures about 3500°K. Therefore, if this model is nearly correct, it seems that the temperature of release is most probably in the vicinity of 3500°K. Specifically, the best all-around fit is obtained at a temperature of about 3500°K and a bursting pressure of over 500 atmospheres. This is probably not too significant in view of the order of magnitude arguments presented. However, the following can be stated. It has been demonstrated that it is possible to ascribe to thermal processes the generation of an ample number of electrons to account for the observed electron cloud. It was assumed that thermodynamic equilibrium was obtained, and that the application of the Saha relationship was justified. The crude models and the liberties exercised for estimating several of the parameters rule out any positive statements. However, if the process of electron generation is ascribed to thermal processes, it is found that reasonable expectation values can be attached to the system discussed.

Table 1 Calculation of the Total Number of Electrons Thermally Produced at Different Assumed Temperatures for 300 and 500 Atmospheres Bursting Pressure of Cannister

T (°K)	% Moles reacted	K_n	Total no. electrons	Total electrons / Total atoms
300 Atmospheres				
1000	24	0·69	$1{\cdot}2\times10^{14}$	$1{\cdot}8\times10^{-11}$
2000	11	$1{\cdot}2\times10^{19}$	$1{\cdot}0\times10^{19}$	$3{\cdot}3\times10^{-6}$
3000	7·1	$4{\cdot}4\times10^{13}$	$4{\cdot}8\times10^{20}$	$2{\cdot}5\times10^{-4}$
5000	4	$5{\cdot}0\times10^{16}$	$1{\cdot}1\times10^{22}$	$1{\cdot}0\times10^{-2}$
500 Atmospheres				
1000	44	0·69	$2{\cdot}7\times10^{14}$	$1{\cdot}4\times10^{-11}$
2000	19	$1{\cdot}2\times10^{19}$	$1{\cdot}3\times10^{19}$	$2{\cdot}5\times10^{-6}$
3000	13	$4{\cdot}4\times10^{13}$	$6{\cdot}3\times10^{20}$	$1{\cdot}8\times10^{-4}$
5000	7·5	$5{\cdot}0\times10^{16}$	$3{\cdot}5\times10^{22}$	$8{\cdot}0\times10^{-3}$

REFERENCES

1. F. F. Marmo, L. M. Aschenbrand and J. Pressman, *Planet. Space Sci.* **1**, 306 (1959).
2. J. F. Bedinger, E. R. Manring and S. N. Ghosh, *J. Geophys. Res.* **63**, 19 (1958).
3. L. M. Aschenbrand, F. F. Marmo and J. Pressman, *Artificial Electron Clouds—VI*, to be published.
4. F. F. Marmo, L. M. Aschenbrand and J. Pressman, *Planet. Space Sci.* **1**, 291 (1959).
5. L. H. Aller, *Astrophysics*. Ronald Press, New York, p. 79 (1953).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 26–32. Printed in Great Britain.

MULTICOLOR PHOTOMETRY OF MARS IN 1958*

G. de VAUCOULEURS

Harvard College Observatory, Cambridge, Massachusetts

(*Received* 26 *June* 1959)

Abstract—Photoelectric observations of Mars in October and November 1958 at the Lowell Observatory give values of the stellar magnitude and integral albedo at five wavelengths from λ 3300 to λ 6900. In the near ultraviolet $3000 < \lambda < 4000$ Mars is "grey" and very dark, with a nearly constant albedo $A = 0{\cdot}046$. Earlier Mount Stromlo and Flagstaff data are also discussed.

1. INTRODUCTION

Photoelectric observations of the stellar magnitude of Mars in five colors were made in October and November 1958 by means of the 21-in. reflector of the Lowell Observatory. The regular photometer designed by H. L. Johnson was used and in addition to the standard *U*, *B*, *V* filters, observations were made through a deep red filter and a silver filter. The effective wave lengths and band widths at half maximum response are listed in Table 1; the values given are for the energy curve of Mars after transmission through the atmosphere.

The spectral response of the 1*P*21 + filter combination was taken from the work of Johnson[1] for the *U*, *B*, *V* filters; the spectral transmission of the silver filter was directly measured; for the deep red combination λ_{eff} was simply derived from the color equation.

The observations at λ 3300 were prompted by a communication from Dr. A. Boggess III reporting that the NRL rocket observations at λ 2700 led to an unexpectedly high value for the albedo, about 0·24[2], in contrast with the very low albedo, about 0·05 near λ 4000[3].

2. OBSERVATIONS

Only three out of 10 scheduled nights were of photometric quality. Much of the time was devoted to the standardization observations necessary to provide a strong tie with the magnitude systems, especially in the *R′* and *U′* bands where no standard magnitudes were available. The magnitudes and colors of the standard stars are listed in Table 2; the number of observations is given in parentheses. For each star the first line gives the standard *U*, *B*, *V* values, the second line the observed values.

Table 1. Filter Combinations

Designation	*U′*	*U*	*B*	*V*	*R′*
λ_{eff}	3300	3600	4550	5550	6900
$1/\lambda$	3·03	2·78	2·20	1·80	1·45
Band-width (Å)	200	550	1100	950	—

The four *A*0 IV,V stars αLyr, γGem, γUMa, γOph were used to fix the zero points of the *U′* and *R′* systems; the four GO V to G5 V stars, ιPer, HR483, 85Peg, κCet whose mean spectral type G2 V and color index $B - V = +0{\cdot}63$ are the same as the sun's[5], were used to determine the other color indices of the sun; four stars in the Hyades, numbers 10, 15, 71 (θ^1Tau) and 72 (θ^2Tau)[4] which were

* The research reported in this paper was supported by the Geophysics Research Directorate, Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF19(604)-3074 with Harvard University.

Table 2. Standard Stars*

Star	Spectr.	V	$B-V$	$U-B$	$U'-V$	$V-R'$
αLyr	*A*0 V	+0·04	0·00	−0·01		
		+0·058	−0·022	+0·04 (3)	−0·010 (1)	+0·015 (2)
γGem	*A*0 IV	+1·93	0·00	+0·03		
		+1·921	+0·012	+0·036 (4)	+0·003 (2)	+0·015 (2)
γUMa	*A*0 V	+2·44	0·00	+0·01		
		+2·414	+0·022	+0·015 (3)	+0·009 (1)	−0·023 (2)
γOph	*A*0 V	+3·75	+0·04	+0·04		
		+3·72	+0·017	+0·004 (2)	+0·125 (1)	—
θ^2Tau	*A*7 III	+3·41	+0·179	+0·132		
		+3·424	+0·181	+0·152 (5)	+0·306 (2)	+0·133 (3)
Hyades 10	—	+7·85	+0·589	+0·100		
		+7·840	+0·593	+0·081 (1)	+0·485 (1)	—
Hyades 15	—	+8·09	+0·658	+0·204		
		+8·052	+0·657	+0·185 (1)	+0·681 (1)	—
ιPer	*G*0 V	+4·04	+0·60	+0·10		
		+4·051	+0·613	+0·094 (6)	+0·591 (1)	+0·494 (3)
HR483	*G*2 V	+4·94	+0·63	+0·12		
		+4·951	+0·614	+0·104 (6)	+0·608 (1)	+0·469 (4)
85Peg	*G*2 V	+5·75	+0·66	+0·04		
		+5·725	+0·662	+0·027 (2)	—	[+0·691] (1)
κCet	*G*5 V	+4·82	+0·68	+0·18		
		+4·830	+0·675	+0·171 (5)	+0·728 (1)	+0·495 (2)
θ^1Tau	*K*0 III	+3·85	+0·995	+0·741		
		+3·850	+0·956	+0·757 (4)	+1·743 (1)	+0·675 (3)

* Number of observations in parenthesis.

close to Mars during the observations, were used as local standards.

Observations were made through a 2 mm = 50″ hole, except on October 28 when the seeing was very poor and a 3 mm = 75″ hole was also used. No systematic difference was noted. The atmospheric absorption coefficients are listed below:

	U'	U	B	V	R'
Q (mag)	0·75	$Q_b+0·26$	Q_b	0·15	0·08

with, as usual, $Q_b = 0·23 - 0·030\ C_g$, where C_g is the color index in the natural system. A plot of the final magnitudes and colors against sec z shows no systematic trend over the whole range $1·0 < \sec z < 3·9$. For sec $z < 3$ the probable errors of one observation are as follows:

	V	$B-V$	$U-B$	$U'-V$	$V-R'$
p.e.	0·02	0·01	0·02	0·03:	0·01:

The relation between $V-R'$ and $B-V$ is linear

$$V - R' = +0·72\,(B-V)$$

The relation between $U'-V$ and $B-V$ is illustrated in Fig. 1. The adopted colors of the sun from the four standard stars of solar type are listed in Table 3; the V magnitude is taken from the work of Stebbins and Kron[8].

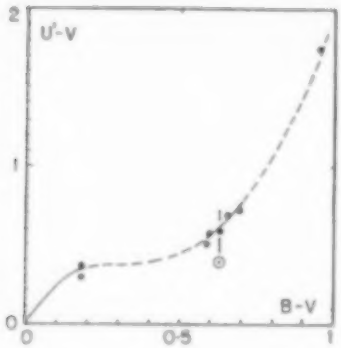


Fig. 1. Relation between U'-V and B-V for the standard stars.

The observed magnitudes and colors of Mars are collected in Table 4.

3. REDUCTION

The magnitude reduced to unit distance from sun and earth ($r\Delta=1$) and to zero phase ($i=0°$) is

$$m_1 = m - 5\log r\Delta - \mu i \qquad (1)$$

The phase coefficient μ varies with wave length as shown by Fig. 2, where the crosses are mean values from the Mount Stromlo photographic spectrophotometry in 1952 and 1954 with double weight for 1954[6, 7] and the dots are from a new reduction of the Flagstaff observations of 1954[8]; the plus is the standard visual value $\mu_v=0{\cdot}015$ mag/deg.[3] The values adopted for the reduction are listed in Table 5; the values for U and U' are extrapolated on the assumption that $\mu=0{\cdot}0185$ mag/deg = const. for $\lambda<0{\cdot}5$.

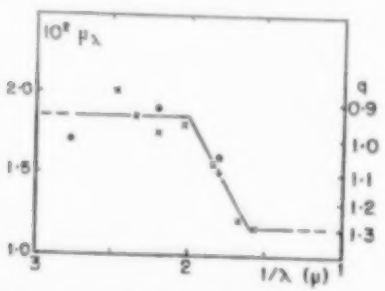


Fig. 2. Phase coefficient of Mars. Crosses: Mount Stromlo 1952, 1954; dots: Flagstaff 1954.

The values in Table 4 so corrected were plotted vs. the longitude of the central meridian ω. The V magnitudes for the same ω are about 0·05 brighter at $i=7°$ to 10° than at $i=21°$ and, since a value $\mu_v\simeq 0{\cdot}02$ is most unlikely, the brightening near opposition may be real; however, fluctuations of the order of $\pm 0{\cdot}05$ mag. are known to occur at all phase angles.[7] Perhaps more significant is the reddening of 0·05 to 0·1 mag indicated by B–V, U–B and U'–V at the smaller phase angles; this may be related to a partial clearing of the blue haze that, according to the observations of Dr. A. G. Wilson (private communication), occurred during November. The effect, if real and significant, may be expected to reach its maximum at exact opposition ($i=0°$), i.e. when the ecliptic latitude of Mars is near 0° at opposition, a circumstance that will occur at the favorable opposition of January 1961.

For the present this effect must be ignored and the linear classical reduction to $i=0°$ according to equation (1) is accepted. The mean magnitudes of Mars at zero phase and unit distance are listed in Table 6 which gives also the derivation of the geometric albedo p and of the spherical (Russell–Bond) albedo $A=pq$ where

$$\log p = 0{\cdot}4\,(M_\odot - m_\beta) - 2\log\sin\sigma_1 \qquad (2)$$

if σ_1 = mean semi-diameter at unit distance and

$$q = 2\int_0^{\pi} \Phi(i)\sin i\,di \qquad (3)$$

Table 3. Magnitude and Colors of Sun

Color	V	B–V	U–B	U'–V	V–R'
m, Δm	−26·73	+0·63	+0·10	+0·63	+0·48
p.e.	0·03	0·01	0·01	0·02	0·02

Table 4. Observed Magnitudes and Colors of Mars

Date (1958)	UT	Long.	V	$B-V$	$U-B$	$U'-V$	$V-R'$
Oct. 23	4:30	95°	−2·034	+1·308	+0·579	+1·975	—
$i=21°\,4$	4:38	98°	−2·027	+1·326	+1·613	—	—
$r=1·453$	7:48	144°	−1·994	+1·287	+0·562	+1·925	—
$\Delta=0·510$	9:08	163°	−1·993	+1·275	+0·542	+1·891	—
	11:56	204°	−2·074	+1·260	+0·644	+1·867	—
	12:04	206°	−2·062	+1·237	+0·626	—	—
	12:41	215°	−2·070	+1·255	+0·641	+1·861	—
	12:48	217°	−2·037	+1·237	+0·616	—	—
Nov. 24	5:15	184°	−2·147	+1·318	+0·710	—	+0·923
$i=6°\,8$	6:14	198°	−2·119	+1·300	+0·700	—	+0·908
$r=1·491$	9:03	240°	−2·050	+1·260	+0·679	—	+0·886
$\Delta=0·510$	11:15	272°	−2·018	+1·271	+0·748	—	+0·854
	11:22	274°	−2·010	+1·310	+0·670	—	—
Nov. 28	2:45	111°	−2·098	+1·323	+0·679	—	+1·033
$i=10°\,2$	2:56	115°	−2·106	+1·334	+0·697	—	—
$r=1·497$	5:16	149°	−2·081	+1·316	+0·686	—	—
$\Delta=0·523$	8:29	196°	−2·122	+1·305	+0·650	+1·924	+1·054
	9:55	218°	−2·082	+1·293	+0·642	+1·883	—

The phase integral which cannot be directly determined since $i<48°$, is estimated by the Russell[9] rule.

$$q \simeq 2·20\,\phi\,(50°) \qquad (3a)$$

or

$$\log q = 0·342 - 20\mu \qquad (3b)$$

(cf. Woolley[6]). For the sake of consistency with earlier determinations we take $\sigma_1 = 4''70$ (cf. de Vaucouleurs[2], p. 280); however, two remarks must be made:

(1) σ_1 is actually the *equatorial* radius, if the optical ellipticity is 0·013 (*loc. cit.* p. 281), the mean radius is 4″67; the corresponding albedo is 1·3 per cent greater than in Table 6;

(2) σ_1 refers to *visual* observations; if the apparent spectral variations of the diameter of Mars are real (*loc. cit.* p. 63) the following *corrections* to A and p may be needed:

Table 4a. Reduced Magnitudes and Colors*

Date (1958)	ω	V	$B-V$	$U-B$	$U'-V$	$V-R'$
Oct. 23	95°–217°	−1·526 (8)	+1·328 (8)	+0·733 (8)	+1·959 (5)	—
Nov. 24	184°–274°	−1·575 (5)	+1·308 (5)	+0·741 (5)	—	+0·894 (4)
Nov. 28	111°–218°	−1·595 (5)	+1·338 (5)	+0·731 (5)	+1·928 (2)	+1·047 (2)
Means		−1·56 (18)	+1·325 (18)	+0·735 (18)	+1·95 (7)	+0·945(6)

* Number of observations in parenthesis.

	U'	U	B	V	R'
$\Delta\sigma/\sigma_v$	+0·03	+0·03	+0·025	0	−0·02
corr. (%)	−6	−6	−5	0	+4

Table 5. Phase Coefficients

Color	V	$B-V$	$U-B$	$U'-V$	$V-R'$
β	+0·015	+0·0035	0·000	+0·0035	+0·0035

Table 6. Reduced Magnitudes and Albedos (1958)

	U'	U	B	V	R'
m_1	+ 0·39	+ 0·495	− 0·235	− 1·56	− 2·505
$m_1-M_\odot$	+26·49	+26·495	+25·865	+25·17	24·705
p	0·049	0·0485	0·087	0·165	0·253
q	0·938?	0·938	0·938	1·100	1·295
$A=pq$	0·046	0·046	0·082	0·181	0·327

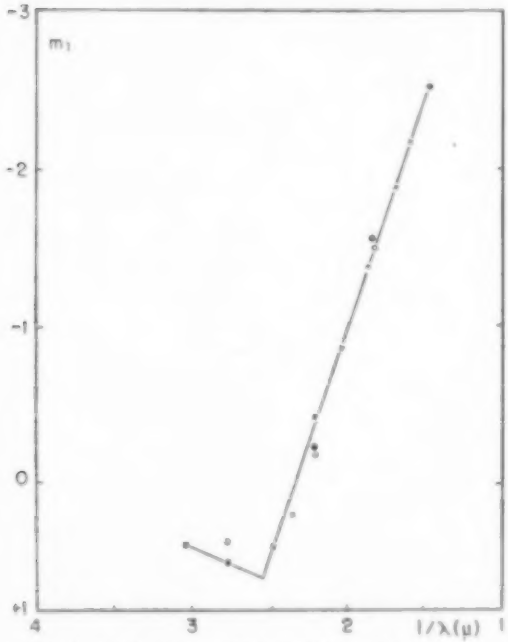


Fig. 3. Reduced magnitudes of Mars ($r\triangle=1$, $i=0$) as a function of $1/\lambda$. Crosses: Mount Stromlo 1952, 1954; circles: Flagstaff 1954; dots: Flagstaff 1958.

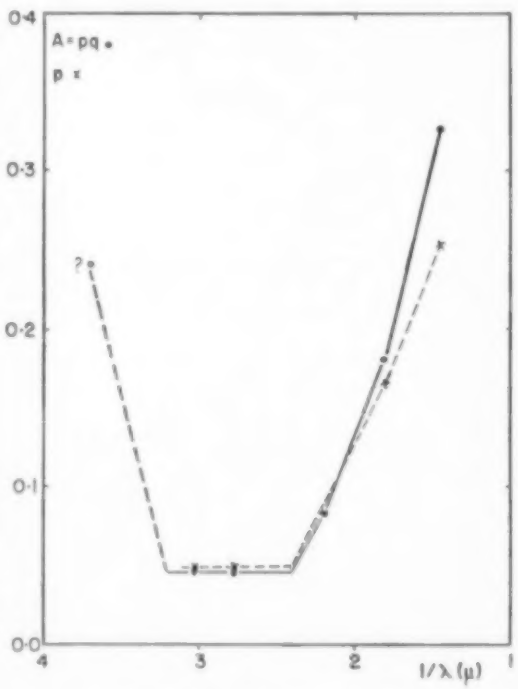


Fig. 4. Spectral albedo of Mars (Flagstaff 1958). Crosses: geometrical albedo p; dots: Russell-Bond albedo $A=pq$; circle: NRL rocket observation.

The values of p and $A=pq$ are plotted in Fig. 4 where the circle marks the NRL result at 2700 Å. The main conclusion is that while Mars is "red" in the visible range $\lambda>4000$ Å, it is "grey" and very dark in the near ultraviolet $3000<\lambda<4000$ Å. If the high albedo at 2700 Å is real, all the increase must take place at $\lambda<3000$ Å and may represent local-

Table 7. Corrected Reduced Magnitudes

λ	Mount Stromlo		Mean	Flagstaff		Mean
	1952	1954		1954	1958	
3600				+0·41	+0·495	+0·45
4050	+0·53	+0·24	+0·40			
4250	+0·29	+0·08	+0·19			
4550	−0·32	−0·53	−0·42	−0·17	−0·235	−0·20
4945	−0·72	−1·04	−0·88			
5430	−1·24	−1·50	−1·37			
5550				−1·49	−1·56	−1·53
5980	−1·86	−1·88	−1·87			
6360	−2·10	−2·21	−2·15			

Note—Comparison with Mount Stromlo spectral data suggests that the Flagstaff values through the *B* filter refer to λ 4500 Å rather than λ 4550 as estimated from the transmission and energy curves.

ized fluorescence in the upper atmosphere of Mars rather than a general albedo increase, due, e.g. to molecular scattering.

4. COMPARISON WITH OTHER DATA

Because of seasonal and accidental fluctuations the albedo of Mars varies slightly with time. A comparison was made with the observations of Woolley *et al.*[6, 7] at Mount Stromlo in 1952 and 1954 and of Johnson and Gardiner[8] at Flagstaff in 1954; the first were by photographic spectrophotometry, the second by photoelectric filter photometry.

These observations were reduced anew with

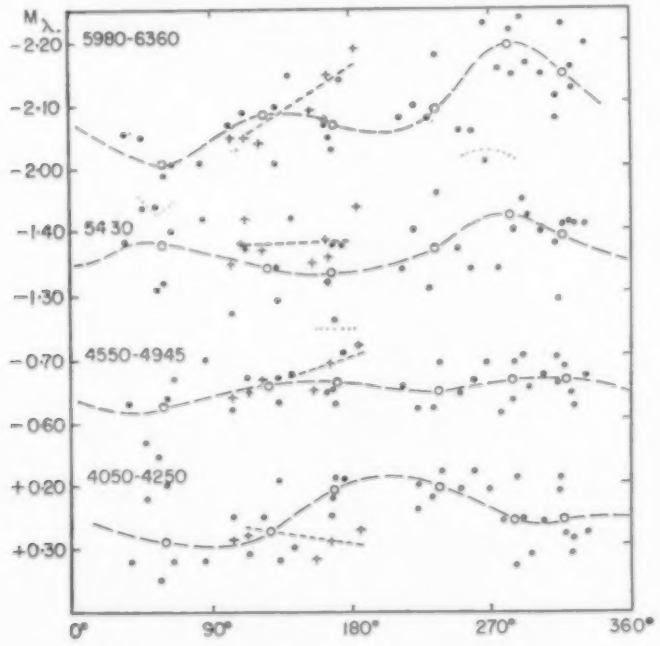


Fig. 5. Magnitude of Mars as a function of central longitude (Mount Stromlo 1954); crosses: (i)<10°.

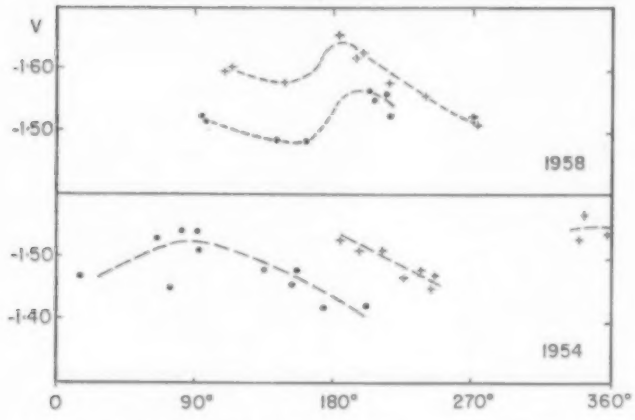


Fig. 6. Magnitude of Mars as a function of central longitude (Flagstaff 1954, 1958). Crosses: (i)<12°.

phase coefficients taken from Fig. 2; in the Mount Stromlo series the original observations had been reduced in terms of $A_{\oplus}-A_{\odot}$ instead of the phase angle i. The corrected values reduced to $i=0°$ are given in Table 7 for the 1954 observations; mean values for λλ 4050–4250, 4550–4945, 5430, 5980–6360 are plotted in Fig. 5 as a function of longitude of central meridian; crosses refer to $|i|<10°$; no systematic difference is indicated. The longitude effect is small, if real it corresponds to maximum light for $\lambda>5000$ Å near longitude $\omega\simeq 280°$; however, this is contradicted by the Flagstaff data for 1954 and 1958 plotted in Fig. 6. Here observations at $|i|<12°$ suggest a brightening near opposition. This should be checked at future oppositions.

Comparison of magnitudes of standard stars indicate that the Mount Stromlo magnitude system requires a zero-point correction of $+0{\cdot}10 \pm 0{\cdot}01$ mag; the corrected reduced magnitudes are listed in Table 7 and plotted in Fig. 3; there is good agreement with the Flagstaff data for both 1954 and 1958. The spectral energy curve of Mars is therefore well established from 3300 to 6900 Å; extensions in both directions should be attempted at future oppositions.

Acknowledgements—I am indebted to the Director and staff of the Lowell Observatory for generously making available the instrumental facilities and in particular to Dr. H. L. Johnson and Dr. W. M. Sinton for their help at various stages.

REFERENCES

1. H. L. JOHNSON, *Ann. Astrophys.* **18**, 292 (1955).
2. A. BOGGESS and L. DUNKELMAN, *Astrophys. J.* **129**, 236 (1959).
3. G. DE VAUCOULEURS, *Physics of the Planet Mars*. Faber and Faber, London (1954).
4. H. L. JOHNSON and C. F. KNUCKLES, *Publ. Astr. Soc. Pacific* **122**, 209 (1955).
5. J. STEBBINS and G. E. KRON, *Astrophys. J.* **126**, 266 (1957).
6. R. V. D. R. WOOLLEY, *Mon. Not. R. Astr. Soc.* **113**, 521 (1953).
7. R. V. D. R. WOOLLEY *et al.*, *Mon. Not. R. Astr. Soc.* **115**, 57 (1955).
8. H. L. JOHNSON and A. J. GARDINER, *Publ. Astr. Soc. Pacific* **67**, 74 (1955). See also correction by A. T. YOUNG, *ibid.* **69**, 568 (1957).
9. H. N. RUSSELL, *Astrophys. J.* **43**, 177 (1916).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 33–38. Printed in Great Britain.

ARCTIC ATMOSPHERIC STRUCTURE TO 250 km*

H. E. LAGOW, R. HOROWITZ and J. AINSWORTH†

U.S. Naval Research Laboratory, Washington 25, D.C.

Abstract—Atmospheric density, pressure and scale height data have been obtained by the U.S. Naval Research Laboratory with Aerobee-Hi rockets fired at Fort Churchill as part of the U.S.-IGY program. Three rockets have been fired at Fort Churchill, 59°N; a fourth will be fired there in October 1958 and a fifth rocket will be fired at White Sands, 33°N in the summer of 1959. The analysis of the data from the first two rockets is nearly complete. The results show: (i) the summer densities from 30 to 70 km at 59°N are approximately 5 to 10 per cent higher than the Rocket Panel values for 33°N, (ii) the winter densities from 25 to 40 km are 10 to 20 per cent lower than corresponding summer values at 59°N; thus they are 5 to 10 per cent lower than the 33°N Rocket Panel values, (iii) the daytime atmospheric density at 200 km at 59°N is higher both in the summer and in the winter than the summer value of $1{\cdot}4\pm0{\cdot}5\times10^{-7}$ g/m³ for 33°N obtained in August 1951. At 59°N and 200 km altitude the summer value is $6{\cdot}7\pm2\times10^{-7}$ g/m³ while the corresponding winter value is $3{\cdot}6+3$, $-1{\cdot}5\times10^{-7}$ g/m³ and (iv) the summer atmospheric pressure at 110 km at 59°N is $(6{\cdot}5\pm0{\cdot}8)\times10^{-5}$ mm Hg and is a factor of 2·5 higher than the corresponding 33°N value.

A tentative result from the third flight has been obtained. This winter night flight gives a density value of $1{\cdot}3\pm0{\cdot}6\times10^{-7}$ g/m³ at 202 km above 59°N. This value is a factor of 3 lower than was obtained at the same latitude on a winter day and is equal to the value obtained on a summer day at 33°N.

1. INTRODUCTION

The Naval research laboratory has flown three of the five Aerobee-Hi rockets which were to be instrumented for atmospheric structure measurements in the IGY Program. These rockets were fired at Ft. Churchill at 1100 hr CST 17 November 1956, 1600 hr CST 29 July 1957, and 1300 hr CST 24 February 1958. This program will be completed with the firing of a fourth Aerobee-Hi rocket at Ft. Churchill in October 1958 and a fifth at White Sands in 1959. Partial results of the first and second rockets were reported[1]* at Toronto, Canada, in September 1957; a final report of the results above 100 km has been submitted for publication[2]. The following paper briefly summarizes the measuring techniques and the main results. Finally, these results are compared to air densities that have been determined by drag measurements on artificial satellites.

2. THEORY OF ROCKET MEASUREMENTS OF ATMOSPHERIC STRUCTURE

It has been demonstrated[3, 4] that when rocket altitude and velocity are known, measurement of pitot-static and stagnation pressure permits a unique determination of the altitude profile of ambient pressure, density and temperature. It has also been shown[3, 5, 6] that at altitudes above 100 km both ambient pressure and density can be measured by a pressure gage mounted on the side of a rocket, provided that the rocket has both a rapid spin and an angle of attack. Here, ambient pressure is obtained by selecting those pressure readings for which the rocket velocity is perpendicular

* July 1958, CSAGI General Assembly, Moscow; published in IGY World Data *A*, *Rocket Report No. 1*, p. 38.

† U.S. Naval Research Laboratory.

to the gage orifice direction, at which time no ram or rarefaction exists at the gage orifice. These ambient pressure points can be selected without knowing the rocket's angle of attack on a rapidly spinning rocket because they occur twice each cycle at exactly one-fourth of a cycle before and after the pressure maximum point. The ambient density at high altitudes is obtained by (*i*) measuring the difference P, between the ram and rarefaction pressures obtained during a single spin cycle, and (*ii*) computing the maximum value of the component of the rocket's velocity which is perpendicular to the plane of the gage orifice.

The detailed theory in references (3), (5) and (6) shows that the density is: (*i*) directly proportional to the pressure change P, (*ii*) inversely proportional to the maximum velocity component perpendicular to the gage orifice, and (*iii*) independent of the ambient temperature. This density measurement requires: (*i*) that the rocket completes a spin cycle at essentially a constant altitude, and (*ii*) that the rocket's angle of attack is known and constant during this spin cycle.

3. INSTRUMENTATION

Fig. 1 shows a Navy Aerobee-Hi rocket. One particular modification made to the rocket for the atmospheric structure experiment is that the tail fins have been canted to impart a spin of 1 to 1·5 c/s. The locations of the pressure orifices are indicated on the pitot-static probe and on the nose cone extension.

Fig. 2 shows one of the cold cathode Philips ionization gages. It should be noted that the gage orifice has a knife edge. The pressure measurement range of the Philips gage is from 10^{-3} to 10^{-7} mm Hg; but with the condenser coupled cathode follower circuit used in the IGY rockets, pressure changes of 10^{-8} mm Hg are measurable.

Fig. 3 shows the 3·5 in. diameter pitot-static probe. Rapid pressure response is obtained by mounting the seven mechanical pressure gages inside the probe and immediately adjacent to the pressure chambers. Two bellows-potentiometer gages are used for measuring stagnation pressures of from 760 mm Hg to 20 mm Hg. Four membrane gages are used to measure stagnation and ambient pressures from 20 mm Hg to 10^{-3} mm Hg. Membrane deflections are measured using a capacitance bridge coupled with an ionization transducer. Sensitivities of as much as 0·2 V per micron Hg are obtained. A fifth membrane gage is condenser coupled to an amplifier and has a sensitivity of 1 V/2 $(10)^{-3}$ mm Hg.

In addition to the pressure sensors each rocket carries: (*i*) a 15-channel PPM/AM telemetering system; (*ii*) an optical aspect system to determine the direction of the rocket's axis relative to the sun and to the horizon; (*iii*) a magnetic aspect system; (*iv*) a radar beacon for trajectory determination; and, (*v*) a DOVAP system for trajectory determination and for telemetry.

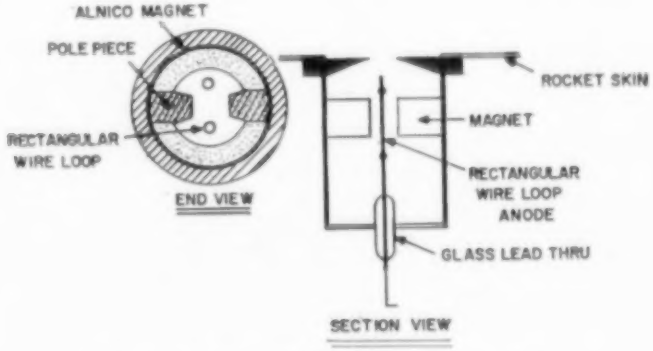


Fig. 2. Cold cathode Philips ionization gage.

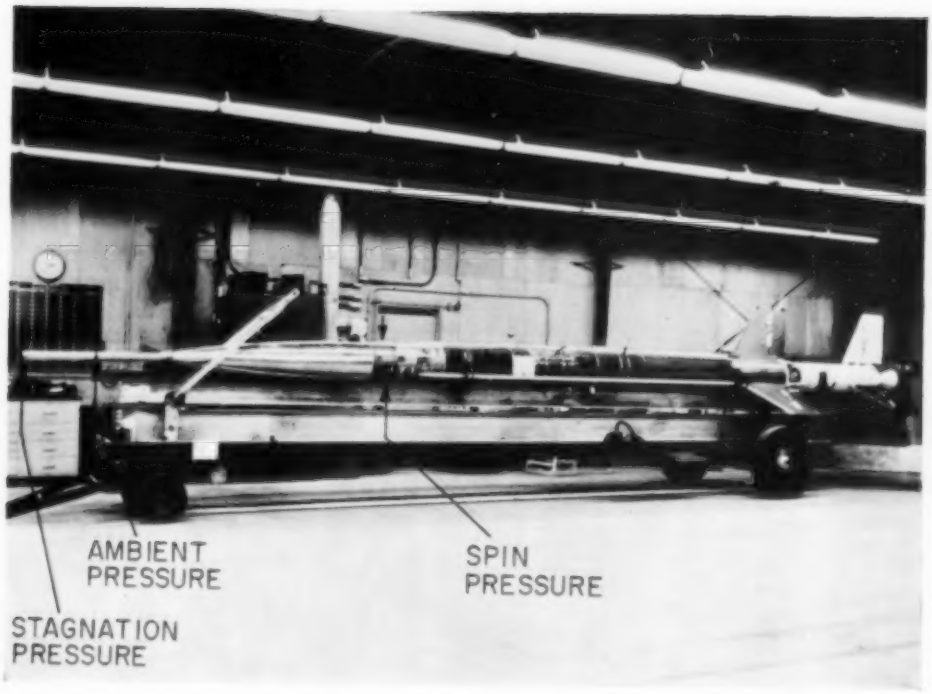


Fig. 1. Aerobee-Hi rocket instrumented for atmospheric structure measurements.

Facing p. 54

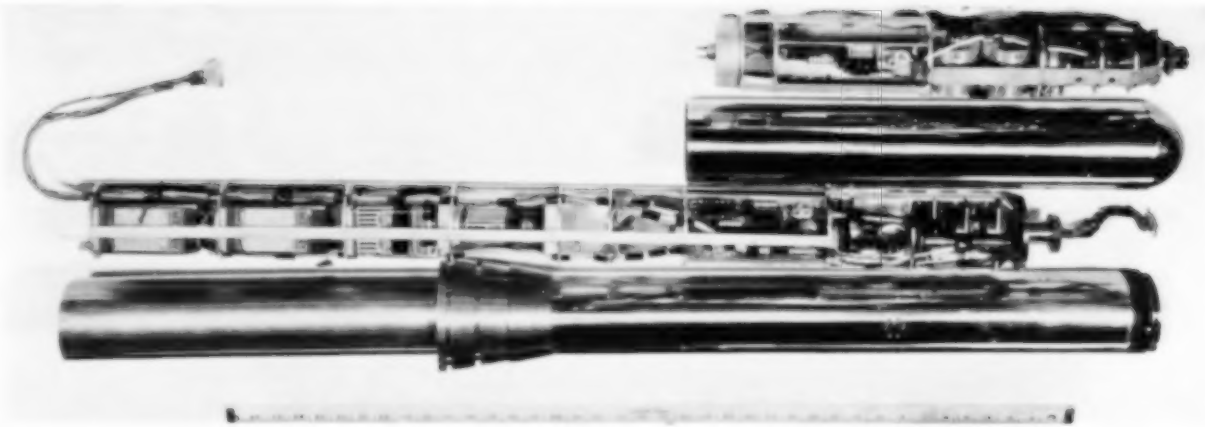


Fig. 3. 3·5 in. pitot-static probe with gages.

4. THE ATMOSPHERE TO 100 km

In Fig. 4, density values are shown as points for both a summer and a winter day at 59°N; the solid curve shows the Rocket Panel data which is an average of all White Sands rocket measurements obtained before 1952. Note that at 25 to 40 km: (*i*) the density points show a seasonal effect of approximately 10 to 20 per cent with summer values being higher; (*ii*) the Rocket Panel densities fall between the summer and winter data. The maximum difference between either the summer or the winter densities and the Rocket Panel value is about 10 per cent. This supports the conclusions regarding the summer densities obtained from 4 Arctic Rockoon[4] flights. Also, note that the summer density points continue at about 10 per cent higher than the Rocket Panel up to 60 km and decrease to the Rocket Panel value at 70 km.

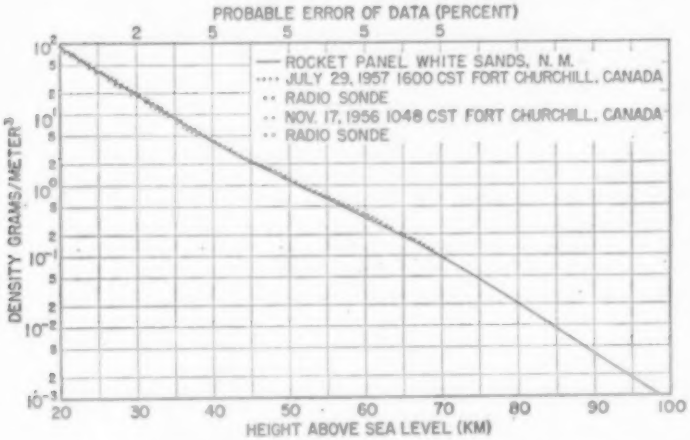

PROBABLE ERROR OF DATA (PERCENT)
2 5 5 5 5
ROCKET PANEL WHITE SANDS, N. M.
JULY 29, 1957 1600 CST FORT CHURCHILL, CANADA
RADIO SONDE
NOV. 17, 1956 1048 CST FORT CHURCHILL, CANADA
RADIO SONDE
DENSITY GRAMS/METER³
HEIGHT ABOVE SEA LEVEL (KM)

Fig. 4. Meaured summer and winter densities below 100 km.

Fig. 5 shows the pressure points for 59°N from 58 to 73 km. They are quite near the solid Rocket Panel curve and have a slope which is approximately the same as the slope of the curve for 74°N.

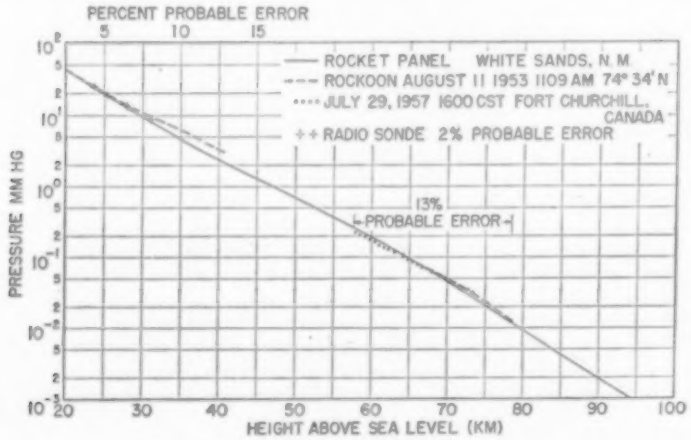

PERCENT PROBABLE ERROR
5 10 15
ROCKET PANEL WHITE SANDS, N. M.
ROCKOON AUGUST 11 1953 1109 AM 74° 34' N
JULY 29, 1957 1600 CST FORT CHURCHILL, CANADA
RADIO SONDE 2% PROBABLE ERROR
13%
PROBABLE ERROR
PRESSURE MM HG
HEIGHT ABOVE SEA LEVEL (KM)

Fig. 5. Measured pressure below 100 km.

5. THE ATMOSPHERE ABOVE 100 km

Fig. 6 shows 59°N summer densities as the solid curve and points. The dashed curve depicts 33°N summer densities[5]. The points on the curve from 125 to 210 km were obtained from two independent Philips gages measuring the pressure modulaton due to rocket spin. The separate measurements were in excellent agreement with each other throughout the flight, and from each gage the same density values were obtained for ascent and descent portions of the trajectory. The density curve was extended down to 100 km by using the pressure curve discussed in the next section and the scale heights derived from both curves. A comparison of these two curves shows a much warmer atmosphere at 59°N and an atmosphere which is more dense than that at White Sands by a factor of 3 at 100 km and a factor of 6 at 200 km. While these large differences can reasonably be ascribed to the heating caused by the long arctic day, to auroral activity and to low energy corpuscular radiations being magnetically focused on the polar regions, one must not ignore the possibility of transient variations in the density. For example, the sunspot number for the 33°N flight of August 1951 was 83 while it was 173 for the 59°N flight of July 1957.

The density data available from the second and third flights are shown in Fig. 9. The winter day value of $3{\cdot}6 \pm 3$, $-1{\cdot}5 \times 10^{-7}$ g/m^3 for 200 km has been published[1]. The winter night point of $1{\cdot}3 \pm 0{\cdot}6 \times 10^{-7}$ g/m^3 for 202 km is from a preliminary analysis of the flight data. Complete processing of the data for this flight will obtain density, pressure and scale height profiles.

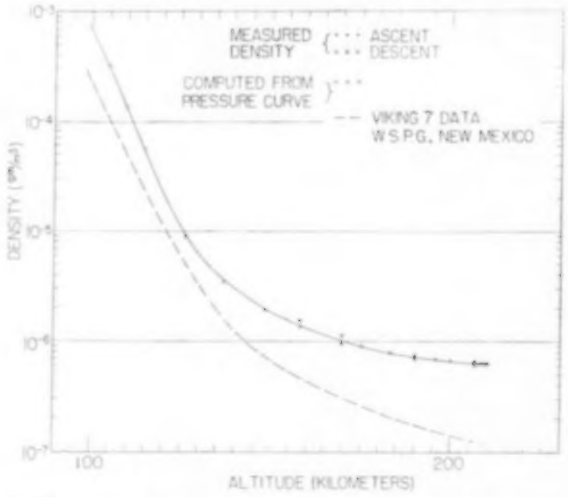


Fig. 6. Summer densities at 59° N. above 100 km.

Pressure data obtained on the summer day flight is shown in Fig. 7. Points above 120 km are computed from the density and scale height data[2]. The dashed curve is for the Viking 7 flight at 33°N, August 1951.

Fig. 8 shows scale heights vs. altitude for the 29 July 1957 IGY firing and the Viking 7 firing. These scale heights, RT/Mg, were derived[2, 5] from the hydrostatic equation and the density and pressure curves of Fig. 6 and 7. To make these calculations it was assumed that the scale height monotonically increases with altitude. The 2 curves each indicate that the scale height gradient is a maximum in the 130 to 150 km region and that the gradient above 180 km is small. A small gradient at these and higher

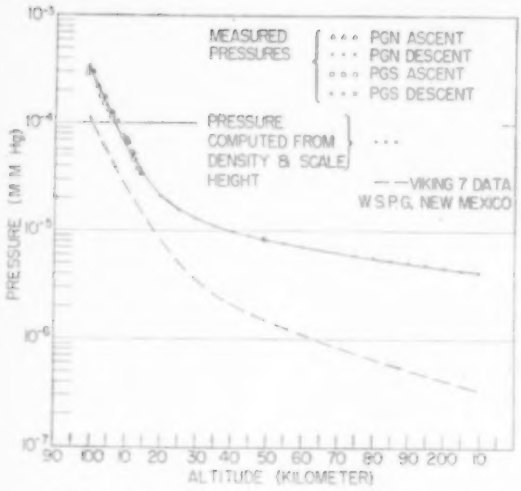


Fig. 7. Summer pressure date at 59° N. above 100 km.

altitudes would be expected from a consideration of thermal conduction[7] in the high atmosphere with low densities if the high atmosphere is being heated by solar radiations.

Fig. 9 shows an extrapolation of the rocket high altitude density measurements out to satellite altitudes by means of the scale height values shown in Fig. 8. Also shown are the winter day and winter night points near 200 km from the rocket measurements at 59°N. All other values are from satellite drag measurements reported by the Smithsonian Astrophysical Observatory[8] with the exception of density points provided by Dr. Jastrow at NRL. Additional points for the U.S.S.R. Satellites 1957_1 and 1957 have not been included because of uncertainties in the mass to area ratio. It is interesting to note: (*i*) the quite close agreement between the 1957_2 points and the rocket curve for 59°N, and (*ii*) that all the satellite density points fall between the 2 rocket

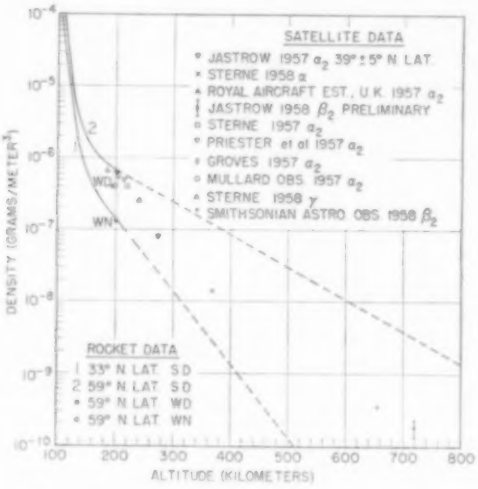


Fig. 9. Extrapolated density and satellite data.

curves. This fact suggests that the average scale heights for the high atmosphere lie between 43 and 95 km and that they are closer to the higher value.

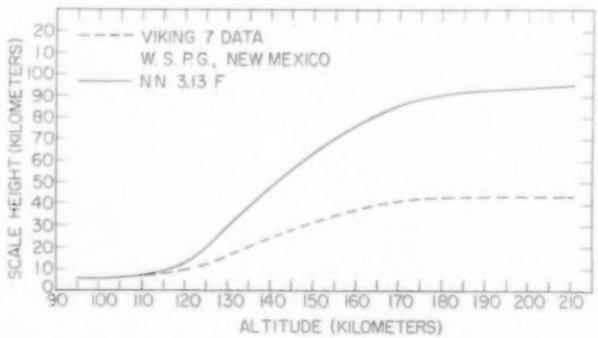


Fig. 8. Summer scale height above 100 km.

6. CONCLUSIONS

The IGY rocket program has demonstrated the existence of large variations in the pressure, density and scale heights of the atmosphere above 100 km. These variations appear to be seasonal, diurnal, and latitudinal with summer values higher than winter, with day values higher than night values, and with Arctic values higher than temperate values. This conclusion is tentative because of the limited number of measurements obtained and because of observed transient phenomena such as sunspots, auroral activity and general circulation, all of which could be significant factors in determining the structure of the high atmosphere.

Finally, it should be noted that these variations can be best studied by using a satellite instrumented to measure spin pressure modulations. The satellite should have a polar orbit with an apogee somewhat greater than 1000 km and a perigee in excess of 160 km. Further rocket measurements will be necessary to study the region below 160 km because of the extremely short lifetime for satellites below this altitude.

Acknowledgements—The authors wish to thank Dr. H. E. Newell, Jr., and Mr. J. W. Townsend, Jr. for their encouragement in this research, and Mr. H. B. Benton for the development and calibration of the electronics instrumentation.

REFERENCES

1. H. E. LaGow, R. Horowitz and J. Ainsworth, *Rocket Measurements of the Arctic Upper Atmosphere. Ann. Geophys. Year*, p. 26–37 (1958).
2. R. Horowitz and H. E. LaGow, *J. Geophys. Res.* **63**, 757 (1958).
3. R. J. Havens, R. T. Koll and H. E. LaGow, *J. Geophys. Res.* **57**, 62 (1952).
4. H. E. LaGow and J. Ainsworth, *J. Geophys. Res.* **61**, **77** (1956).
5. R. Horowitz and H. E. LaGow, *J. Geophys. Res.* **62**, **57** (1957).
6. R. Horowitz and D. Kleitman, *U.S. Naval Research Laboratory Report* No. 4246 (1953).
7. D. R. Bates, *Proc. Phys. Soc.*, **B64**, 805 (1951).
8. *IGY WDC A—Rockets and Satellites—Satellite Report* No. 4 (1958).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 39–48. Printed in Great Britain.

THE FALL-OUT OF METEORIC IRON PARTICLES

E. A. KREIKEN

Ankara University Astronomical Institute, Ankara, Turkey

(*Received* 24 *February* 1959. *Revised* 24 *June* 1959)

Abstract—The present paper is a survey of the attempts, which during recent years have been made at the Astronomical Institute of the Ankara University:

(1) to determine the daily amounts of the fall-out of iron particles of meteoric origin;

(2) to study the seasonal variations of the fall-out; and

(3) to compare the curve representing these seasonal variations of the fall-out with the one representing the seasonal variations in meteoric activity.

The results obtained up till now are critically discussed and possibilities for improving the methods of observation are indicated.

1. INTRODUCTION

The curve with which that of the fall-out has to be compared, e.g. the curve for the seasonal variations in meteoric activity, is itself not very well determined. A composite diagram of the curves obtained by several observers using different methods has been given by Lovell[1]. In the majority of these curves a pronounced maximum in meteoric activity is indicated for August, while generally a secondary maximum seems to be present around October–November.

But while as far as these main features are concerned there is a general agreement between the different curves, there are considerable differences in details. Secondary waves, which appear to be clearly indicated in one of the curves, do not show up at all in the other curves and vice versa. Also, with at least two of the curves given in Lovell's diagram, the October–November maximum in meteoric activity appears to be more important than the August one. With all other curves the reverse is the case.

This is not as strange as may seem at first sight, because it is now generally known that the apparently simple problem of an accurate and nonselective determination of the meteoric activity offers many intricate observational and theoretical difficulties. The problem of determining the amounts of fall-out of meteoric material should be much simpler because all difficulties due to directional selection and to apparent magnitude are eliminated automatically. But actually, from the experience at this Institute, it appears that to obtain data on the fall-out which numerically are reliable is not too easy.

So far all attempts have been limited to obtain data about the fall-out of iron particles, because these are the only ones which can easily be collected and can be separated from the normal amounts of dust of purely terrestial origin. That a steady fall-out of meteoric particles must be expected to occur, follows from theoretical considerations. For this the reader is referred to the studies of Hoppe[2], Oleak[3] and others. As appears from theory, after the impact of the initial meteoric mass on the upper layers of the earth's atmosphere and after the subsequent losses of mass and energy, there remains a particle of small residual mass and of small and uniform residual velocity. The size of both the residual mass and of the residual velocity depends on the initial conditions, such as initial mass, initial velocity and angle of incidence. As is discussed in Section 3 of this paper, it is in this respect that the observations

on the fall-out give some direct and positive results from which it is evident that in the lower strata of the atmosphere the gravitational and frictional forces almost completely balance. It takes the particles at least 48 hr to travel through the final 2 km of the atmosphere. Consequently, one must not expect the relation between the curves of meteoric activity and of the fall-out to be a very direct one. According to theory, the limiting small value of the residual velocity is already attained at a great distance (at least 20 km) above the surface of the earth. Although the evidence about the small size of the residual velocity actually applies to the final 2 km only, we may regard this as a proof that the theoretical results are substantially correct for a much larger part of the orbit. We may then take it that the residual meteoric particles may, for considerable intervals of time, remain semi-suspended in the air and, with very small but still individually different velocities, only gradually will settle down on the surface of the earth. While during this long time interval the particles of different residual mass will travel with different velocities, they will at the same time be subject to the influence of large and small scale currents and eddies in the earth's atmosphere. This means that by the time they settle down on the earth, they have completely lost their identities. Particles, originating from different meteor streams and from sporadic meteors, will have been thoroughly intermixed.

Thus, what we must expect to find when the curves for the seasonal variations of the meteoric activity and the fall-out are compared is that the latter will be a replica of the first, but with the details smoothed out. Also, there may be a time delay between the two curves.

In the attempts to evaluate the daily amounts of the fall-out, two different methods have been used. The first method is based on direct counts of the numbers of iron particles which in the course of one day settle down on a surface of a given area. This method and the results obtained so far are discussed in Section 2, 3, 4 and 5.

The second method is less direct and is based on estimates on the amount of darkening that particles, collected from the same surface, cause to appear on a white piece of plastic. The results are discussed in Sections 6 and 7.

As yet neither method has given results which are entirely satisfactory. Ultimately, the curve of the fall-out will have to be compared with the curve of meteoric activity based on day and night observations, e.g. those based on radar observations made at the same latitude. (This because there are reasons to suspect that the fall-out may not be identical in all latitudes.) However, in the determination of both the curve of the fall-out and of the meteoric activity, there remain so many doubtful points that at present such details can be neglected.

Perhaps some additional information could be obtained if measurements were made at various sites: the fall-out of iron particles cannot always safely be ascribed to cosmic origin. In at least one series of the observations, which will be described, the influence of heavily industrialized areas shows up. From the series of observations made at Ankara it appears further that the numerical results are strongly affected by climatological circumstances. Days with rainfall cause large fluctuations to appear in the observed curve. The nature of these fluctuations is such that even when mean values of several days are used, the resulting curve is quite unreliable. Therefore, attempts to determine the shape of the curve on the fall-out should be restricted to favourable sites. These are sites outside industrialized urban districts and in regions with little rainfall. In this respect the situation at Ankara is good, because there hardly is any industry while long and uninterrupted stretches of dry weather occur in all seasons of the year. On the other hand, here also large accidental errors show up in the determination of the daily amounts of fall-out. But the observations could be continued over such long periods of time that it was possible to trace the origin of most of the errors, and it is hoped that with future observations they can—partly at least—be avoided. It still will be valuable if observations are repeated at another site, for it is known that impurities can be carried over long

distances through the air. Moreover, the effect of different latitudes might be of relevance.

2. COUNTS OF PARTICLES

Results of continued sets of observations in which the numbers of particles in the fall-out were daily counted were published by Kizilirmak[4] and Süslü[5].

Here only a slightly different version of the composite curve given by Süslü is reproduced (Fig. 1). This curve gives the seasonal variations of the numbers of particles in the fall-out. The scale is indicated on the left hand side of the figure, while the months of the year are indicated at the bottom. Black dots indicate the results obtained by Kizilirmak during the year 1954, open circles those obtained in 1955 by Süslü. The black circles and dots indicate normal points which generally represent the mean of ten consecutive observations. For reasons which will be explained, there had to be excluded not only the observations made on days of rainfall, but also those on the first and second day after rainfall. So for several normal points less than ten individual observations were available. In the figure this appears from the sizes of the discs and the dots which were taken proportional to the number of observations, the largest size corresponding to ten observations. The general shape of the two curves is identical. From a deep minimum around February, the curves rise with irregular fluctuations until a September maximum. A secondary maximum seems to be present around October–November, but the minimum around September–October needs further confirmation as do all of the secondary fluctuations in the curves. This reservation should not only be made on the general grounds stated in Section 1, but also because:

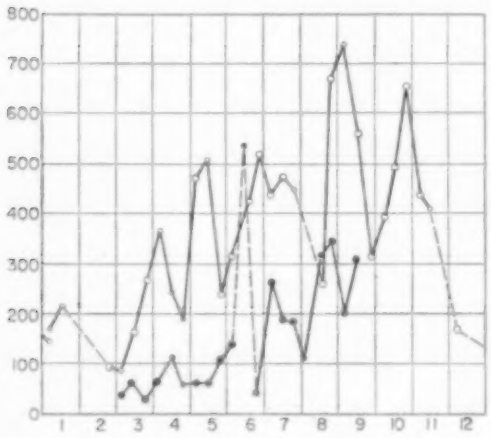


Fig. 1. Seasonal variations of the numbers of particles in the fall-out. Upper curve by Süslü, lower curve by Kizilirmak. Ordinates are observed numbers per day on surface of 0·8 m².

(1) the standard error in the determination of the individual observations is large, and
(2) there obviously is a strong personal bias in the way the numbers are counted.

As far as the standard error is concerned, the following remarks must be made. The counts were continued during the year 1956 in such a way that the numbers of particles were determined by both observers independently. In a subsequent paper, Kizilirmak[4] has reported on these further observations. He has determined the size of the standard error in two different ways. In the first place a smooth curve was drawn through the normal points, based on ten observations each, and the scatter of the individual points around this smooth curve was evaluated. Let N_i be the numbers of particles as obtained from the individual observations, N_0 those read from the normal curve. Then the error affecting the ratio N_i/N_0 can be obtained from the usual relation:

$$\sigma = \mp \sqrt{[\log (N_i/N_0)^2]/n}$$

For the two observers Kizilirmak finds the numerical values $\sigma_k = \mp 0{\cdot}29$ and $\sigma_s = \mp 0{\cdot}25$. When the values obtained by the two different observers are compared in a similar way, the error in the ratio N_{ik}/N_{is} is found. Its numerical value appears to be $\sigma = \mp 0{\cdot}38$ which would correspond to an error of $\mp 0{\cdot}28$ for each individual observer. The size of these standard errors is unsatisfactorily large. An error of $\mp 0{\cdot}38$ in the logarithm of the ratio N_{ik}/N_{is} corresponds to an error of $(2\text{–}4)^{\pm 1}$ in

the ratio itself. This means that when two different observers under identical circumstances count the number of particles on one and the same plate, it is quite possible that the one observer finds a number which is either 2–4 times larger or 2–4 times smaller than the number obtained by the other observer. For such a simple problem these differences seem to be astonishingly large and it may be worth while to probe into the reasons for this. Before proceeding to do so, it should be remarked that there is some other point with the numerical value of σ which needs clarification:

When for each observer the value of σ is determined from a direct comparison of the two sets of values N_i, the resulting value of σ corresponds to a pure observational error. Thus, $\sigma^2=\sigma_0^2$. On the other hand, when the numerical value of σ is derived from the scatter of the individual observations around a smoothed mean curve, the resulting error σ represents the combined effects of two sources of error. Of these the first is the pure error of observation and the second is due to the true accidental variations in the daily amounts of the fall-out. Thus, $\sigma^2=\sigma_0^2+\sigma_r^2$. We should therefore expect the latter value to be the larger one. Actually the two values which Kizilirmak gives are of almost equal magnitude. It would be premature to conclude from this that the accidental variations in the amounts of the daily fall-out are small. This is not so because the size of the σ values is too large to allow any definite conclusion, and because a direct comparison of the two sets of values N_i may give a value of σ which is too large. As remarked before, Fig. 1 gives rather strong evidence that there is considerable systematic difference between the two sets of observations, and this has not been taken into consideration in the values $\sigma^2=\sigma_0^2+\sigma_r^2$ given by Kizilirmak, mainly because this systematic difference seems not to be quite constant.

It seems probable that both the personal bias and the large size of the standard errors are due to the distribution curve of the particle sizes. In his extensive study on meteoric dust, Hasegawa[6] gives the distribution curve by particle size obtained by different observers, mainly in Japan, together with the distribution curve as deducted by Millman[7] from theoretical consideration. All curves have one feature in common: With decreasing values of the diameters there is an exceedingly rapid increase of the numbers of particles. Several of the observed curves show a sharp drop of the numbers of particles beyond the limit 5μ, but this must be ascribed to the fact that here we are beyond the limit of completeness. It is easy to understand that the extreme steepness of the distribution curve of the particle sizes must have affected the series of observations made at Ankara. When two different observers count the numbers of particles on the same plate, the limit of completeness down to which they count is not necessarily the same, and even with one and the same observer, due to the influence of personal disposition, fatigue, etc., daily fluctuations are unavoidable. Though the difference between the two limits may be small, it still will result in a considerable deviation in the observed numbers of N_i. Presumably this explains the large systematic difference in Fig. 1 between the results of two different observers. From the above it would appear that both the large size of the standard error and the pronounced personal influence of the individual observer in the future possibly can be avoided when the diameter of the particles is also taken into account.

3. CLIMATOLOGICAL INFLUENCE ON THE COUNTED NUMBERS OF PARTICLES

Especially during and shortly after periods of rainfall, the numbers of particles which are counted vary in a highly irregular way. At the request of the author, Kizilirmak has collected all data which were available for cases when either one day of rainfall or a short period of rainfall were both preceded and succeeded by continued stretches of dry weather. In the mean the total numbers of particles before and after rainfall should be equal, but this does not imply that equal numbers will be observed. During rainfall the volume of air between the rain-producing layers of a cloud and the surface of the earth will be swept clean of particles.

So immediately after rainfall no particles will settle down on the surface of the earth. They will only reappear after a time interval equal to the time the particles need to travel the distance from the clouds to the earth. Altogether Kizilirmak found eleven cases for which this could be tested. For these cases he has tabulated the observed numbers 4, 3, 2, 1 and 0 days before and 1, 2, 3 and 4 days after rainfall. Instead of reproducing his tabulated numbers here, his results have been represented graphically in Fig. 2. As expected, on the days of rainfall itself hardly any particles are observed. It is interesting to observe that when the particles reappear, their numbers do not immediately return to normal. After a sharp initial rise, the gradient of the curve indicating the numbers decreases and finally the curve merges asymptotically into a horizontal one. Therefore several interesting conclusions can be drawn from Fig. 2, though as yet merely qualitative ones.

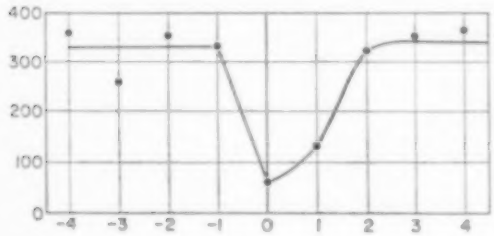


Fig. 2. Influence of rain on the numbers of particles.

From meteorological data it appears that the height of the rain-producing clouds in the mean was 2 km. Fig. 2 proves that for the swiftest moving particles it takes less than 1 day to travel through these final 2 km, while for the slower moving particles this may take up to 3 days and more. From these data a certain mean residual velocity can be derived, and it is satisfactory that this is of the same order of magnitude as predicted by theory. For obvious reasons, one expects the differences in the residual velocities to be related to the initial condition, e.g. the initial velocity, the angle of incidence and especially the initial mass. The particles of large residual mass also should be those with large residual velocity. Unfortunately, Fig. 2 gives no information about this point. Also the curve in Fig. 2 is a mean obtained from observations, distributed over the whole year. However, in the different parts of the year the mass distribution curve of the particles is not necessarily a constant one. In this case also, the conclusion is that more detailed observations are needed in which the size of the particles is taken into account.

The first aim should be to improve the determination of the shape of the ascending branch of the curve in Fig. 2: The time intervals of one day each are rather too long. The numbers will have to be counted after 6, 12, 18, etc., hr after the end of rainfall; and while the numbers of particles are counted, their size must be estimated at the same time.

In another respect the information given by Fig. 2 is quite definite. It proves that during a period of rainfall and for a few days afterwards, the numbers of particles in the fall-out are abnormal and but remotely related to the true numbers. So whenever the seasonal variations in the fall-out are studied by counting the numbers of particles at any site, all observations made during a few days after either continued or intermittent periods of rainfall must be excluded. Conversely from direct counts, results can only be expected if the observations are carried out at a site where during any part of the year long and uninterrupted stretches of dry weather occur. Violent storms are extremely rare at Ankara, so that our observations do not contain any direct information about this point, but it is obvious that all observations carried out under such unfavourable circumstances are better discarded.

4. THE INFLUENCE OF INDUSTRIAL ACTIVITIES ON THE COUNTED NUMBERS

During the year 1956 a simultaneous set of observations on the fall-out were carried out at the Heidelberg Observatory (Germany) by Kizilirmak and by Süslü at Ankara. It was hoped that in this way sufficient material for intracomparison could be obtained. The method used was that of the counts of particles. During that year weather conditions at Heidel-

berg were unfavourable, so that continued sets of observations could hardly be obtained and no curve for the seasonal variations could be obtained. So the series of observations carried out at Heidelberg need only be considered briefly. Still one result of more general interest was obtained. The area to the northwest of Heidelberg is heavily industrialized so that there was a possibility to test the influence of industrial activities on the fall-out. The observed daily numbers were arranged into groups, each group corresponding to a certain direction of the prevailing wind that day. When for each group mean values were computed, it appeared that these mean values systematically depended on the direction of the wind. A maximum occurred when the wind was from the northwest, i.e. from the direction of the industrialized area. Also, as observed through the microscope, in the fall-out many spherical dark particles occurred which in general outward appearance seemed quite different from those usually observed. For details the reader is referred to the original paper by Kizilirmak[3]. For obvious reasons a similar check was applied to the Ankara series of observations. In the latter case no influence of the wind direction on the observed numbers was detectable. So it would seem that the Ankara set of observations is free from the influence of industrial activity. It is known that particles of very small size can be carried through the air over very long distances, and so this last conclusion cannot as yet be safely put in an entirely definite form. On the other hand this simultaneous set of observations contains definite proof that observations on the fall-out should not be carried out in industrialized areas. But this is a point that is almost self-evident and it is highly improbable that any observer would select such a site to start observations on the fall-out.

5. THE CURVE FOR THE SEASONAL VARIATIONS

The discussions in the previous sections have given some insight as to the meaning of the numbers which are observed. Therefore it is worth while briefly to reconsider the evidence contained in Fig. 1 about the seasonal variations in the curves on the fall-out. The large systematic difference between the two curves may now be taken to indicate the difference between the limits of completeness which the two different observers have attained. But the size of this systematic difference makes it obvious that any attempt to give a numerical interpretation of the curves would be premature. Also, there is the disadvantage that the great size of the systematic difference makes it impossible to combine the results obtained by different observers into one set. As yet the curves have to be studied separately. The question may be raised as to how far a qualitative interpretation is already possible. Some general trends in the observed numbers seem to be well established.

From a minimum in March the fall-out in the mean increases until a certain maximum is reached and next decreases again. The exact time of the maximum is of interest. From Fig. 1 it would appear that this maximum at the earliest is in September. Compared to the curve of the meteoric activity this is a delay of 1 month. Such a delay is within the possibilities. It would be interesting to have this time difference fixed in a more accurate way, but for this neither the curves for the meteoric activity nor those on the fall-out are sufficiently determined. It is to be observed that the curves on the fall-out are mean curves valid for a mixture of particles of many different sizes, while it is by no means certain that for particles of unequal sizes identical curves for seasonal variations are valid. In this case also, before drawing any definite conclusions, the results of further observations must be awaited in which the sizes of the particles have been taken into account. As to the many jagged peaks and irregularities which appear both on the ascending and descending branches of the curves in Fig. 1, as yet little can be said. The appearance of some peaks roughly coincides with the occurrence of one of the well known day or night time meteoric streams. At different instances meteoric streams are observed throughout the course of the whole year. Therefore, as long as there is no 1:1 corres-

pondence, such coincidences must be considered as accidental. This does not mean that a possible relation between a given peak in the curve on the fall-out and one of the meteoric streams must *a priori* be denied. But on the ground of the present material such a direct relationship cannot be established.

Summarising, in all questions we have taken up in the preceding sections, we were confronted by the necessity of having further observations in which the numbers of particles were not only counted but in which also their sizes are considered. It is evident that such a program, however necessary, will be elaborate and time consuming. Consequently, first the possibilities of another method were considered, from which it was hoped to obtain quick and accurate results. The first experimental stages of this alternate method have been completed and are described in the subsequent sections. As yet the results have not come up to expectations, but from a critical examination it would appear that they can be considerably further improved.

6. THE METHOD OFF "STEP" VALUES

This method was first tried by Kizilirmak[6]. Instead of a bar magnet an electromagnet of constant field strength, 500 G, was used. To one end of the bar magnet a piece of white plastic was attached. The iron particles in the fall-out were collected in the same way as before by sweeping the pole of the magnet back and forth over the surface of the glass plate on which the particles settle down. In this way on the plastic a faint black ring is formed.

It is evident that the intensity of the ring will directly be related to the numbers of particles. In first approximation this relation can be expected to be a logarithmic one. The rings were obtained day after day and from these a sequence was selected of increasing darkness. These rings of different standard intensity were numbered S_1, S_2, ... S_{10}, so that S_1 corresponds to minimum intensity while S_{10} is the one of largest intensity. Once the standard sequence S has been selected, the intensity of a ring obtained at any particular day can be fitted into the standard sequence by simple eye estimate, while decimal fractions can be estimated. The method therefore closely resembles Argelander's method of "Stufenschätzungen" which very accurately gives results with the variable stars. An attempt was made to adapt this method to less favourable weather conditions. Instead of the large glass plate, two wide rectangular trays were used in which the particles were allowed to settle down during periods of 24 hr. Of these trays one was filled with water, the other was kept dry. On a day of rainfall following a stretch of dry weather, the rain will carry along the particles in the air and these particles will be deposited in the water-filled tray. So it is to be expected that immediately after rainfall the dark ring obtained from the wet tray will have too large an intensity. This reasoning seems to be confirmed by the observations. It may be assumed that at different altitudes the net flux of the particles is a constant. Undoubtedly the flux density will vary with the size of the particles, but if the rain continues after the first day, the numbers of particles deposited by the rain in the wet tray will, for each particle size, exactly correspond to the net daily flux of particles of that peculiar size. Therefore—at least theoretically—the intensity of the ring obtained from the wet tray should be normal from the first day on that rain has started and should remain so during continued periods of rainy weather. Whether this prediction is actually fulfilled for the time being is doubtful. At Ankara continued stretches of rainy weather are rare and only occur during the winter months. Whether the S values then obtained are normal or not, can only be decided when they are compared with S values obtained during spells of dry weather, though obviously during a spell of dry weather in the corresponding season of a different year. But during the winter months reliable S values for dry weather periods are difficult to obtain so that for a direct comparison the material is inadequate. Therefore in this paper all evidence as to the intensities S as obtained from the wet tray have entirely been neglected. An additional reason for this is that unexpectedly the standard error in the

determination of the daily values S proves so large as to be unsatisfactory. It can immediately be determined by comparing the S values obtained every day and during a long period of time from two trays with equal surface (both dry). The resulting numerical value of the standard error proves to be $\sigma_s = \mp 0{\cdot}07$ which roughly corresponds to an error of $\sigma = \mp 0{\cdot}21$ in the logarithmic ratio of N_2/N_1. This error is a purely observational one. Usually, with Argelander's method of step values, the internal error of the estimates proves to be small. Therefore, the large size of the standard error which turns up in this case must be taken to indicate either that there are considerable local irregularities in the density of the fall-out or that the method used needs further improvement. Though the possibility of local irregularities cannot be denied, they can hardly be expected to be of the magnitude as suggested by this numerical value of σ. Hence, it must be expected that the large error mainly arises from the way the method has been executed.

The first point which must be critically examined is the way in which the relation between S and the corresponding values N was established. For all days on which both S and N had been observed, the values S were plotted against the corresponding values log N. Through the points obtained in this way, a smooth curve was drawn. This curve represents the mean relation between S and log N. It is not necessary to reproduce this curve here. It is sufficient to state that it is a non linear one. For an analytical representation some equation of the form

$$S = A + B \log N + C (\log N)^2 + \ldots .$$

would have to be used and the numerical value of the coefficient C would be relatively large. Both with S and log N only the intermediate values are frequent.

As the standard errors in both the determination of S and log N are large, this means that the curve remains badly determined at both ends and that only the intermediate part is acceptable. For future work it seems preferable to set up a new standard scale of S values in such a way that a purely linear relation between the values S and log N is assured in advance. If such a curve could be set up, only its zero point would have to be fixed. It should be possible to establish such a purely linear scale. A further question which must be considered is whether the constancy of the standard scale of S values which was used is sufficiently guaranteed. The standard rings $S_1 \ldots . S_{10}$ which were selected, were placed between glass plates and kept in a dark place. They were always handled carefully when the estimates of the daily values S were made, but it still is dubious whether the intensities $S_1 \ldots . S_{10}$ will remain a constant. The samples consist of loose particles which have been knit more or less together under action of the magnet. There seems to be the danger that gradually the aggregate of particles will loosen up and the rings will fade. Also deterioration under influence of atmospheric circumstances cannot be discarded without more consideration. Consequently the scale which has been used until now must be considered as being unreliable. The only way out for future work seems to fix the scale photographically. This implies that also photographic reproductions will have to be made of the rings which are collected daily. As long as care is taken that this is done under standardized conditions there is no objection to this. Standardization of the scale will not be sufficient unless at the same time means are found substantially to decrease the size of the standard error of the individual values of S.

As described before, the rings are formed on the plastic when the magnet is swept back and forth over the bottom of the trays. When the electromagnet is swept over the surface by hand, the iron particles will mainly be collected on the side of the forward motion of the magnet. Hence, when the magnet is swept back and forth over the surface, the ring thus formed will have a tendency to have too large a density at two places opposite each other while the intermediate regions will have a lower density. Also when the magnet is swept by hand over the surface, the distance between the

bottom of the tray and the end of the magnet will not always remain a constant. Further objections against the way the alternate method has been executed are:

(1) The magnet will not always remain in a vertical position; this may cause additional irregularities in the outward appearance of the ring.

(2) When the surface is swept over by hand, we cannot be sure that all different areas have been covered with equal care; there may occur variations from one area to the other and also from one day to the other.

It would seem that the large errors which turned up in the estimated values of S can largely be explained from the various causes mentioned above. This implies that the method of execution urgently needs improvement.

7. SEASONAL VARIATIONS IN THE S CURVE

In his paper Kizilirmak[4] has given a curve representing the seasonal variations of the observed values of S for the year 1957. With a view to the many objections raised in the preceding sections, his curve must be considered as a preliminary one needing further improvement. For this reason it is not reproduced here. A short description will suffice, while it is interesting to notice that its general shape is of the form which should be expected.

From a minimum around March the curve rises in an irregular way to a maximum around October and next decreases again. Both the amplitude of the major wave and those of the irregularities on the ascending and descending branch of the curve are smaller than those which appear in Fig. 1. This is natural because the values of S are roughly proportional to the logarithms of the numbers and therefore, relative to the curves representing numbers, the S curve will be flattened. In this case also with a view to the large size of the standard error affecting the S value, no attention will be paid to the secondary variations. It is satisfactory that in having one pronounced maximum and one pronounced minimum the general shape of the S curve seems to correspond to those based on the direct counts of the numbers of particles in the fall-out. The time of the minima seems to be the same with all curves. Relative to the curves in Fig. 1, with the S curve there seems to be a time lag in the occurrence of the maximum. As yet with all curves the exact occurrence of the maximum is but indifferently determined and subject to doubt. So this time lag must be accepted with great reserve and needs further confirmation. Some caution is needed because the S curve and those on the fall-out are not necessarily identical. The curves in Fig. 1 refer to particles down to a certain limiting size of the diameter. It may be presumed, that the estimates of the intensity S refer to all particles including those which can no longer directly be counted. Theory predicts that the numbers of these infinitesimal particles are very large and consequently their contribution to the total intensity S may not be negligible. So only if in the lower layers of the air we always have to deal with the same mixture of particles, e.g. if the distribution curve of the particle sizes remains a constant, a one-to-one relation between the numbers N and the intensities S can be expected. When large-scale irregularities occur in the flux density of the particles colliding with the upper layers of the atmosphere, there is no longer any certainty that the mixture of particles in the lower layers will actually remain constant. So certain relative shifts between the S and N curves cannot *a priori* be considered to be impossible. This point needs further investigation, but such investigations must be postponed until sufficient accurate material has become available and the shapes of the S and N curves have been determined in a more accurate way. From these considerations it is evident that the seasonal distribution curve S has a restricted meaning. At any time the value S will indicate the total intensity due to all particles together and without respect to the distribution curve for the diameters. The distribution curves corresponding to equal values of S may be different. In this respect the information given even by an improved S curve is bound to be less complete than that given by direct counts of particles in which the sizes of

the particles are also taken into consideration. But on the other hand, because the S values also include the very small particles, they will give a valuable extension of the observations towards these smaller particles, so that their special behaviour can be studied.

It must be observed that the estimates of the S values were introduced in order to have a method which would give quick and accurate results and which could be continued even during periods of unfavourable weather. From an inspection of the results hitherto obtained, it has appeared that the degree of accuracy which is desirable has not been attained.

8. THE TOTAL MASS OF THE FALL-OUT

Nothwithstanding their large numbers, the combined weight of all particles which daily set on a limited surface is small. However, for the two curves which are given in Fig. 1, a single smoothed curve can be provisionally substituted. Without any danger this smoothed curve can be extrapolated so as to cover the whole year. Let this provisional curve be indicated by $f(t)$. It is evident that when the combined weight w is determined of all particles in the fall-out during a certain period, say from t_1 to t_2, the total weight W of the fall-out in the course of a year is

$$W = w \, x \int_0^{365} f(t)\,dt \; : \; \int_{t_1}^{t_2} f(t)\,dt$$

The values of the integrals can be read from the smooth curve. Conversely, once W has been determined, this relation enables us to find the weight of the fall-out during any subperiod. It is obvious that the choice of the initial instants t_1 and t_2 must be made in such a way that during this interval there is uninterrupted dry weather.

Kizilirmak[4] has applied this method to his observations. During a period of 95 days ($t_1 = 29$ June 1956; $t_2 = 4$ October 1956) the iron particles were allowed to settle on a plate with a surface of 1·28 m^2. At the end of this period the iron particles were collected on a small glass plate attached to the end of a bar magnet. For w the value $(1{\cdot}6 \pm 0{\cdot}1)$ mg was obtained. If it is assumed that the fall-out is uniform over the whole surface of the earth and if for this surface the value $S = 5 \times 10^{14}$ m^2 is adopted, for the total fall-out in the course of a year the value

$$W = (3{\cdot}1 \pm 0{\cdot}2) \times 10^9 \text{ kg}/S$$

is obtained. As yet this number needs further confirmation. Bearing in mind that only the iron particles in the fall-out have been collected, this provisional value would seem slightly on the large side.

It would seem that from the series of measurements on the fall-out which have been carried out at Ankara some significant results have been obtained: they are significant because they represent first results. But the accuracy of the measurements proved to be less than was hoped for. On the other hand, we believe that we have been able to trace the reasons for the large accidental errors which turn up in both methods. If our reasoning was correct it should be possible in the future substantially to reduce the size of the standard errors.

REFERENCES

1. A. C. B. Lovell, *Meteor Astronomy*, p. 97. Clarendon Press, Oxford (1954).
2. J. Hoppe, *Astr. Nach.* **262**, 12 (1937); *Mitt. Sternw. Jena* **15** (1954), *ibid.* **23** (1956).
3. H. Oleak, *Mitt. Sternw. Jena* **25** (1956).
4. A. Kizilirmak, *Comm. Astr. Inst. Ankara* **33** (1954), *ibid.* **33** (1959).
5. R. Süslü, *Comm. Astr. Inst. Ankara* **22** (1956).
6. I. Hasegawa, *Smithson. Astrophys. Obs., Ref. Man.* **1** (1959).
7. P. M. Millman, *J. R. Astr. Soc. Canada* **46**, 79 (1952).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 49–55. Printed in Great Britain.

OUTLINE OF A THEORY OF MAGNETIC SEPARATION OF AURORAL PARTICLES AND THE ORIGIN OF THE S_D FIELD

J. E. SHAW

Ionospheric Prediction Service, Department of the Interior, Miller's Point, Sydney, N.S.W.

(*Received* 20 *June* 1959)

Abstract—The diurnal variations of occurrence of two types of sporadic *E*-ionisation at Macquarie Island (geomagnetic latitude 61·1°S) are identified with the diurnal variations of the flux of incoming protons and electrons. It is shown that protons and electrons which have been accelerated under the control of the geomagnetic field, will be separated by magnetic forces resulting from their motion along the curved geomagnetic lines of force.

The resulting pattern of precipitation in the atmosphere explains qualitatively the regular diurnal features of the occurrence of sporadic *E*-ionisation. An extension of the theory provides a basis for interpretation of the observed development of a visual auroral display, with particular reference to the mode of formation of visual auroral forms. The theory explains the origin of the field responsible for the S_D magnetic variations and is in agreement with the spiral occurrence of magnetic disturbance variations observed in the north polar regions by Meek[1].

1. INTRODUCTION

Recent reports of observations carried out during auroral disturbance by Romick and Elvey[2], Galperin[3] and others, show the importance of hydrogen ion bombardment in the formation of a display. Their observations show that the maximum of hydrogen emission does not correspond in time with maximum visual auroral activity, and that radio echoes may also be received when there is little or no visual auroral activity.

Ionospheric observations carried out at Macquarie Island (54·5°S, 159·0°E) at geomagnetic latitude 61·1°S, are in agreement with these observations. The development of an auroral display as observed from one site, has a time sequence which, when averaged over a number of displays, produces a consistent diurnal pattern. The analysis of echoes from sporadic *E*-ionisation shows that there is a corresponding diurnal pattern in the ionisation present in the *E*-region. In this paper it has been assumed that two identifiable types of sporadic ionisation result from entry into the *E*-region of hydrogen ions and electrons after being accelerated by a process controlled by the earth's magnetic field.

The resulting pattern of particle precipitation explains the regular features of the development of an auroral display, and shows how auroral observations may be interpreted to provide information about the accelerating process. The theory provides a basis for the study of the origin of visual auroral forms and demonstrates the likely importance of the S_D and dynamo electric fields in the development of these forms. The part played by the dynamo field in the control of radio wave absorption in the auroral zone is discussed.

2. THE DATA

Published results of ionospheric[4] and magnetic[5] observations made at Macquarie Island have been used in the analysis. At the beginning of the International Geophysical Year, classification[6] of the sporadic *E*-types observed in the routine ionosonde records was commenced. In Fig. 1(a), (b), are shown the percentage occurrence curves for 10 months from July 1957 to May 1958 inclusive, for the prin-

cipal nocturnal types *f*, *a* and *r*. Type *l* has been included with type *f* because the seasonal variation of sunset and sunrise times makes the distributions for these two types overlap.

Figure 1(b) also shows the percentage occurrence of complete absorption derived from the hourly observations for 4 equinoctial months during the period September 1951–March 1953. The form of this distribution is maintained for all equinoctial and summer months at Macquarie Island and the times of occurrence of the main features of the distribution are independent of mean solar activity. The distribution has therefore been used for comparison with sporadic *E*-occurrences in 1957–1958.

The use of *B* as a quantitative measure of absorption has numerous difficulties. In the subsequent discussion, however, the only part of the distribution referred to, occurs between the hours 2000–0500, 157·5° E.M.T. During this period the principal experimental effect is the gradual decrease in *F*-region critical frequency, an effect which can be expected to produce a gradual increase in the occurrence of *B* without any quantitative change in the amount of absorption. More rapid variations during this period can be attributed to variations in the absorption of the reflected wave.

The magnetic data, Figs. 3 and 4, are obtained from the hourly values of mean horizontal intensity and the daily time of occurrence of maximum disturbance in horizontal intensity measured during the period 1952–1954. The times of occurrence of the principal features in these curves are also independent of mean solar activity, and they have therefore been used for comparison with more recent data on sporadic *E*-occurrences without prejudicing the conclusions drawn.

3. DISCUSSION OF THE OBSERVATIONS

From Fig. 1(a) it is evident that the occurrence of sporadic ionisation at Macquarie Island is distributed fairly symmetrically about midnight. This fact is confirmed by the diurnal distribution of sporadic *E*-percentage occurrence of all types over a period of 7 years at this station. When the percentages of occurrence of the different types are separated, two subsidiary peaks *F*, *F'* and *A*, *A'*, are present at 2130 and 0230 hr. The relative percentages of occurrence of *f* and *a* types are consistent with the observations of a number of authors, namely that *a* type sporadic *E* is associated with enhanced auroral and magnetic activity. The two peaks, *F* and *F'*, could well be derived from the occurrence of *a* type sporadic ionisation because the transition from *f* type

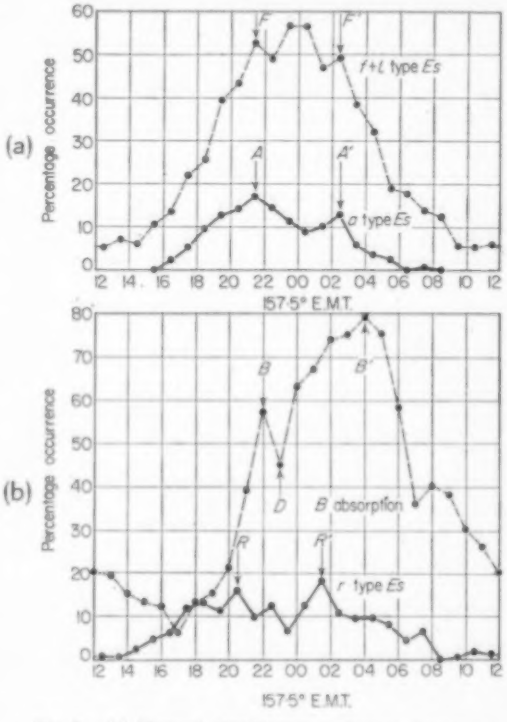


Fig. 1. (a) Diurnal distribution of hourly percentage occurrence of sporadic *E*– types at Macquarie Island. (i) *f* + *l* type E_s : ●— — —●. (ii) *a* type E_s: ●———●.

(b) Diurnal distribution of hourly percentage occurrence of high absorption and sporadic *E* at Macquarie Island. (i) high absorption: ●— — —●. (ii) *r* type E_s: ●———●.

to *a* type is not clearly defined on many ionograms.

Figure 1(b) shows the hourly distribution of occurrence of *r* type sporadic ionisation and *B* (complete absorption). *r* type ionisation is

distinctive in form, displaying group retardation near the critical frequency and thereby showing the characteristics of reflection from a layer of considerable depth, with a maximum of electron density at some distance above the bottom of the layer.

Two principal peaks of occurrence, R and R', are evident for r type ionisation at 2030 and 0130 hr. It may be noted that these peaks are displaced by 1 hr from the peaks of occurrence in the f and a type distributions, and that the percentage occurrence of r type ionisation at these peaks is approximately equal to that of the a type ionisation.

Working at College, Alaska, Romick and Elvey[2] have found that the zenithal maximum of intensity of $H\beta$ radiation almost invariably precedes the maximum in intensity of the 4709 Å line of N_2^+ by times greater than 1 hr. $H\beta$ radiation in the zenith is normally associated with quiet forms of visible aurora in the lower sky, while the 4709 Å line is associated with the more active forms of aurora in the zenith. Further they find no evidence of the dispersal of $H\beta$ radiation over the whole sky at the onset of more active auroral forms.

Little and Leinbach[7] have tentatively correlated high absorption at vertical incidence with pulsating aurora, while Major[8], after observing at Macquarie Island, has stated: "There is a closer resemblance between the nocturnal variations of intensity of the aurora and frequency of occurrence of complete absorption than there is between frequency of occurrence of pulsating and flaming auroras and complete absorption."

When Fig. 1(b) is examined it is found that the form of the equinoctial absorption curve, between the hr 2000–0500, follows very closely the form of the average hourly intensity of the 4709 Å line of N_2^+ at College. The form of the r type sporadic ionisation curve and its position in time relative to the absorption curve are similar to the average hourly intensity of $H\beta$ radiation at College.

4. MAGNETIC SEPARATION OF AURORAL PARTICLES

The following assumptions are made to explain the time variations of sporadic ionisation and auroral occurrence.

(a) The E-region reflections which are classified as r type E_s, are derived from ionisation caused directly by the entry of hydrogen ions or protons into the region, and the time variations of r occurrence represent the time variations of the hydrogen ion flux.

(b) The E-region reflections, which are classified as a type E_s, are derived from ionisation caused directly by the entry of electrons into the region, and the time variations of a occurrence represent the time variations of auroral electron flux.

These assumptions are consistent with the observed height distribution of ionisation in the two forms. r type E_s is characterised by the presence of group retardation which indicates an extended region of ionisation above the lowest height of occurrence of a type E_s. This would be expected to result from the difference in penetration of protons and electrons whose energies are approximately equal.

(c) The particles are accelerated by a process such as that proposed by Cole[9] so that their mean energies are approximately equal and the major component of particle velocity is along the earth's magnetic field.

(d) The accelerating process reaches field lines which terminate in the lowest magnetic latitudes on the magnetic midnight meridian.

The last assumption is supported by the observations of a number of authors of the apparent expansion of the auroral zone into lower magnetic latitudes near magnetic midnight. At Macquarie Island, Jacka[10] has shown that there is a deviation of the latitude of occurrence of homogeneous arcs from a linear regression on K_p (geomagnetic planetary disturbance index) which has its greatest low latitude extent near magnetic midnight.

The process of magnetic separation is now explained with reference to the model diagrammatically shown in Fig. 2. The model is based on a centered dipole field whose axis coincides with the geographic axis of rotation. Particles are accelerated in the accelerating region and spiral along the magnetic field lines. If these

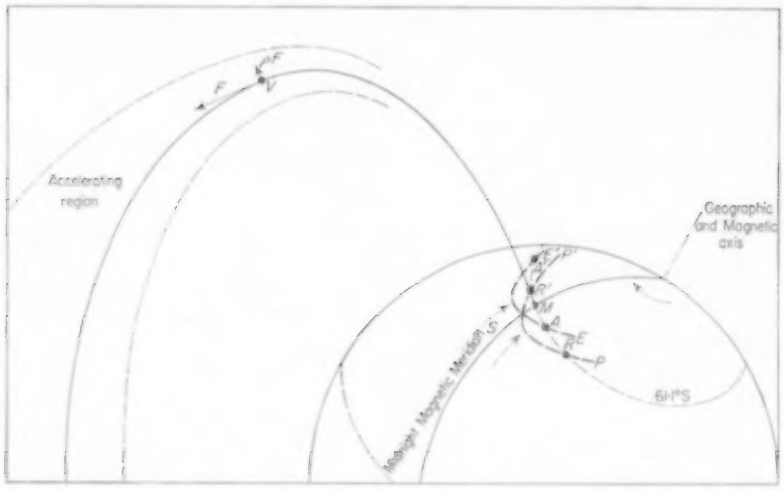


Fig. 2. Model of the process of magnetic separation of auroral particles as observed at Macquarie Island, with coincident geographic and magnetic axes.

lines were straight, assumption (d) would require that the maximum flux of both protons and electrons would reach its lowest latitude on the midnight meridian. The curvature of the magnetic field lines produces an effective component of magnetic force perpendicular to the field line for particles travelling along it. If at the point V, the tangential magnetic vector is F, then for a particle spiralling along the field line, there is an effective magnetic vector (ρF) along the radius of curvature of the field line, where ρ is the curvature of the field line at V.

This radial force deflects protons and electrons in opposite directions at right angles to the plane of the field line. On arrival at the E-region, the particles which were originally accelerated along the midnight meridian field line are displaced in space, so that the low latitude extent of the maximum particle flux appears to the observer at M, on the midnight meridian, to be displaced in time. EE' and PP' are curves representing the locus of the spatial distribution of maximum particle flux relative to the midnight meridian for electrons and protons, respectively. As the observer at M moves in time with respect to the midnight meridian he observes the flux maxima at points R, A, R', A', in turn. These are identified with the peaks of occurrence in the a type and r type distributions of Fig. 1.

The spatial separation, S, of the maximum low latitude extent of proton and electron precipitation may be inferred from the separation in time of the observed peaks of occurrence R, A and R', A', provided that the form of the loci PP' and EE' is identical.

5. DISCUSSION OF DEDUCTIONS FROM THE THEORY

If the theory is established, some deductions may be made which have an important bearing on the interpretation of observations of auroral and sporadic E-occurrence. These deductions will be discussed below.

(a) Observations made at separate stations on the same magnetic latitude may provide information about the accelerating process for protons and electrons which originated in the same part of the accelerating region.

(b) The spatial separation of protons and electrons will be a function of their energies, with the greatest separation for particles of lowest energy. The apparent separation in

time of events observed during a disturbance from one station, should provide information on the energies attained by the particles and hence the intensity of the accelerating process.

(c) If, from the observations of Romick and Elvey[2] and others, the maximum proton precipitation can be said to occur in lower magnetic latitudes than the quiet auroral forms, it may be deduced that these auroral forms occur between the loci $P\,P'$ and $E\,E'$. The occurrence of the quiet forms should depend on the electric field between the two loci of precipitation and the energy spectrum of the incoming accelerated particles. The subsequent break-up of the quiet forms into an active display corresponds to the arrival of the observer at the region of maximum electron precipitation.

(d) The occurrence of a second maximum of quiet auroral display during one night's observations will require the presence of protons on the locus $P\,P'$ and will depend on the latitude distribution of electron flux.

(e) If the loci $E\,E'$ and $P\,P'$ are continued into the daytime hemisphere and it is assumed that at magnetic noon the accelerating process penetrates only to magnetic field lines terminating in higher magnetic latitudes, then the distribution of auroral occurrence will very closely approximate that described by Meek[1] for magnetic disturbance in the northern auroral zone. This correspondence with the spiral distribution discussed by Meek is enhanced if the energies acquired by the auroral particles are dependent on their position relative to the midnight magnetic meridian. If these energies decrease as the noon magnetic meridian is approached, then the spatial separation of protons and electrons is increased and the resemblance to Meek's spiral distribution is enhanced.

(f) If the auroral arcs are formed in the region between the loci of maximum electron and proton flux, and their orientation is dependent on the direction of the electric field between these two regions, a diurnal variation in the orientation should be observed from one station. Currie and Jones[11] have reported such a variation.

(g) The origin of the electric field which causes the S_D magnetic variations in high latitudes is in the polarisation field maintained by the magnetic separation of the maxima of proton and electron flux. The energy for maintenance of the field is derived from the energy of the particle acceleration process in the accelerating region. This deduction is at variance with the theory proposed by Martyn[12] which places the origin of the polarisation field in the equatorial ring current. Nagata and Fukushima[13] attribute the main disturbance field to the enhancement of the dynamo S_q variations by the entry of additional ionisation into the auroral regions. If Figs. 3 and 4 are referred to, it is evident that the influx of ionisation at Macquarie Island does enhance the night-time dynamo variations, but the theory does not explain the large disturbance increases in the horizontal component of magnetic force which begin at about 1500 hr. Nor does it explain the absence of a seasonal variation in time of onset of the S_D disturbance, whose magnitude is comparable with the seasonal phase variation of the S_q field.

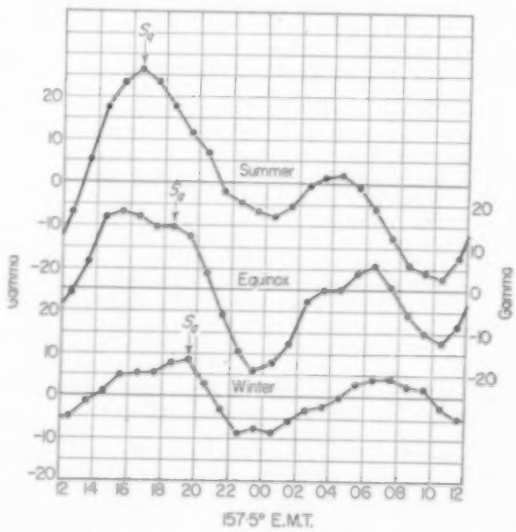


Fig. 3. Diurnal distribution of hourly mean variations of the horizontal component (H), on the 5 international quiet days, for 5 months of each season at Macquarie Island, 1952-1954.

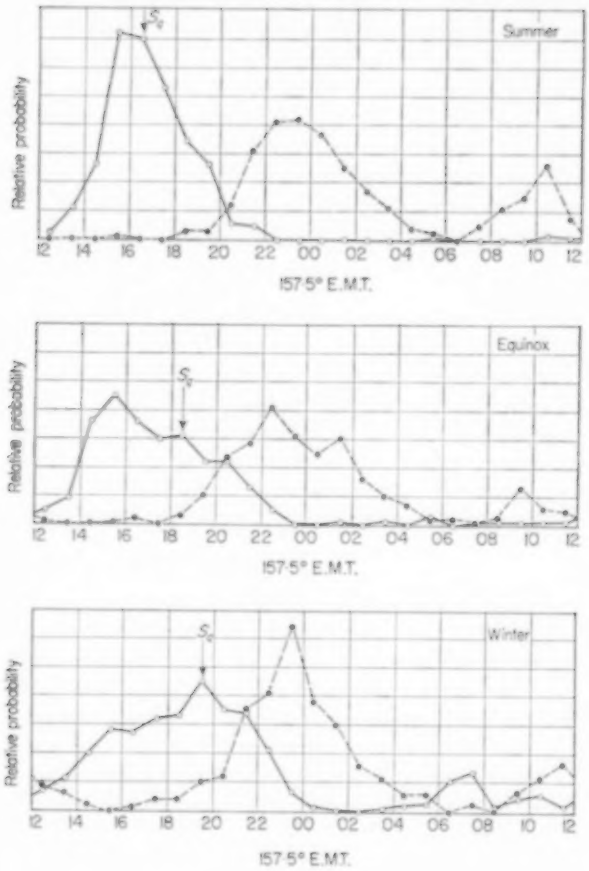


Fig. 4. Relative probability (arbitrary units) that the daily maximum variation in horizontal intensity (H) will occur in the hour shown, for all days of 5 months in each season at Macquarie Island, 1952-1954. Daily maximum in H: O———O. Daily minimum in H: ●— — —●

(h) The similarity between the time variations of occurrence of high absorption at Macquarie Island (Fig. 1(b)), and the time variations of intensity of the 4709 Å line of N_2^+ observed by Romick and Elvey, has been referred to above. Shaw[14] has discussed the evidence that the absorption measured by waves reflected from the ionosphere at vertical incidence is dependent on the current intensity in the E-region and the consequent vertical drift of ionisation.

In Fig. 3 it is seen that the eastward S_q current and therefore the northward dynamo electric field, is approaching a maximum at 0400 hr for the summer and equinoctial seasons. The field is in the right phase to cause an increase in absorption, a process which would be enhanced by the influx of auroral electrons.

Since there is a parallel increase in the 4709 Å auroral line at this time, after the maximum of electron precipitation A' has passed, it appears likely that the form and intensity of the active aurorae are controlled both by the distribution and flux of auroral electrons and the intensity of the dynamo field.

Acknowledgments—The discussed ionospheric, magnetic and auroral observations have been made at Macquarie Island using facilities provided by the Australian National Antarctic Research Expedition of the Department of External Affairs. The assistance of Mr. B. Tierney in the analysis of the magnetic data is acknowledged.

REFERENCES

1. J. H. Meek, *J. Atmos. Terr. Phys.* **6**, 313 (1955).
2. G. J. Romick and C. T. Elvey, *J. Atmos. Terr. Phys.* **12**, 283 (1958).
3. G. I. Galperin, *Planet. Space Sci.* **1**, 57 (1959).
4. *Ionospheric Prediction Service, Commonw. of Aust., Dept. Interior*, Ser. D (1950-1958).
5. *Bureau of Mineral Resources, Geology and Geophysics, Commonw. of Aust., Dept. Nat. Develop.*:
   P. M. McGregor, *Rep.* No. 27 (1956);
   P. B. Tenni and J. A. Brooks, *Rep.* No. 31 (1956);
   C. S. Robertson, *Rep.* No. 35 (1957).
6. J. A. Thomas and E. K. Smith, *J. Atmos. Terr. Phys.* **13**, 314 (1959).
7. C. G. Little and H. Leinbach, *Proc. Inst. Radio Engrs.* **46**, 334 (1958).
8. G. Major, *Aust. J. Phys.* **7**, 471 (1954).
9. K. D. Cole, *Nature, Lond.* **183**, 738 (1959).
10. F. Jacka, *Aust. J. Phys.* **6**, 219 (1953).
11. B. W. Currie and C. K. Jones, *Terr. Mag.* **46**, 269 (1941).
12. D. F. Martyn, *Nature, Lond.* **162**, 92 (1951).
13. T. Nagata and N. Fukushima, *Rep. Ionosph. Res. Japan*, **6**, 85 (1952).
14. J. E. Shaw, *Planet. Space Sci.*, to be published.

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 56–59. Printed in Great Britain.

A THEORY OF THE ORIGIN AND GEOMAGNETIC CONTROL OF TWO TYPES OF HIGH LATITUDE E_S

J. E. SHAW

Ionospheric Prediction Service, Department of the Interior, Miller's Point, Sydney, N.S.W.

(*Received* 20 *June* 1959)

Abstract—Observations of night-time E_s occurrence from four stations near the southern auroral zone are analysed. A major peak in the diurnal curves for the percentage hourly occurrence of *f* type E_s is identified with the peak of occurrence of Auroral Belt type E_s discussed by Penndorf and Coroniti for the northern auroral zone.

It is found that the probable low lattitude extent of this peak is approximately 50° geomagnetic.

A theory is proposed which attributes the occurrence of Auroral Belt type E_s and Thule type E_s to two separate processes, both of which involve the entry of electrons into the ionosphere under the control of the geomagnetic field.

The theory shows that Auroral Belt type E_s may be caused by electrons from a region of trapped particles the extent of which can be defined by the extent in geomagnetic latitude of the night-time peak in *f* type E_s occurrence. Thule type E_s may be caused by an influx of electrons into the ionsphere from a stream of particles approaching from the direction of the sun.

The deflections of the peak of occurrence of *f* type E_s from the midnight magnetic meridian, and of the Thule type E_s from the "geomagnetic noon" meridian, are a function of the mean energy of the electrons and the curvature and intensity of the lines of force in the geomagnetic field.

1. INTRODUCTION

Penndorf and Coroniti[1] have analysed the occurrence of Polar E_s in the northern polar regions and have found two basic types together with a mixed type. The characteristics which were used to define these types are quoted below:

(*a*) "The Auroral Belt type, for stations between 63° and 73° geomagnetic latitude, shows a strong night maximum all through the year and no appreciable change with solar activity."

(*b*) "The Thule type, for stations north of 73° geomagnetic latitude, shows a single maximum in summer around 'geomagnetic noon'. A winter maximum may occur in some years, with no pronounced daily maximum."

(*c*) "The Mixed type shows both types in a more or less mixed form. During some months, the Thule type prevails; in others, the Auroral Belt type prevails. Such behaviour occurs at Baker Lake."

These authors found that the peak of occurrence of Auroral Belt type E_s moves towards later hours in the night for stations in lower geomagnetic latitudes. They concluded that, in view of the close relationship between E_s and magnetic activity especially in the auroral belt, E_s results from the direct or indirect influence of corpuscular radiation on the night-time E_s along the auroral belt and on E_s in general for the arctic region closer to the magnetic pole.

Shaw[2] has proposed a mechanism in which the distribution of two types of E_s, *a* and *r*, is determined by the distribution of the influx of accelerated particles near the visual auroral zone. In this paper, E_s observations from some stations near the southern auroral zone have been further analysed. It is found that the

results of the analysis are consistent with those obtained by Penndorf and Coroniti for Auroral Belt type E_s. The probable low latitude occurrence of night-time E_s resulting from particle bombardment is extended and a theory is proposed which attributes the occurrence of Auroral Belt type E_s and Thule type E_s to two distinct processes.

The theory shows that both Auroral Belt type E_s and Thule type E_s resulted principally from the influx of negatively charged particles. Those causing Auroral Belt type E_s are derived from the disturbance of a region of trapped particles which encircles the Earth in longitude, while those causing Thule type E_s are guided into the region of highest geomagnetic latitudes by the lines of force of the Earth's magnetic field, and are derived from a stream of particles which approaches the Earth from the direction of the sun.

2. THE DATA

Published hourly occurrences of E_s types[3] have been used for four stations: Macquarie Island (*MI*)[4], Campbell Island (*CI*)[5], Hobart (*H*)[6] and Christchurch (*CHCH*)[7], for the period July 1957 to November 1958. In the calculation of the hourly percentage occurrence of f type E_s from the hourly data, the report of a lost record or of complete absorption has been accepted as no observation.

Penndorf and Coroniti have restricted their data by using only those occurrences for which $f_0E_s > 5$ Mc/s. In the analysis reported here, the only restriction placed on the data is that imposed by the method of scaling ionograms. This restriction arises from the fact that, for almost all hours, only the most prominent and consistent type of E_s present is scaled. For night-time E_s in the auroral zone, it is assumed that this method of scaling will give a clear representation of the diurnal variations of the several factors which may be responsible for the occurrence of E_s. The factors which control the intensity of E_s need not necessarily have the same diurnal distribution as those which control its occurrence.

The results of the analysis are shown in Fig. 1, where the diagrams giving the diurnal distribution of percentage hourly occurrence are arranged in decreasing geomagnetic latitude and approximately local time. The data for *MI*, Fig. 1(a), and *CI*, Fig. 1(b), cover all months in the period used. For *H*, Fig. 1(c) and *CHCH*, Fig. 1(d), six equinoctial months have been chosen. The last two stations are in geomagnetic latitudes less than the auroral zone and it was considered that the equinoctial curves would be more likely to enhance the occurrence of E_s resulting from the influx of particles.

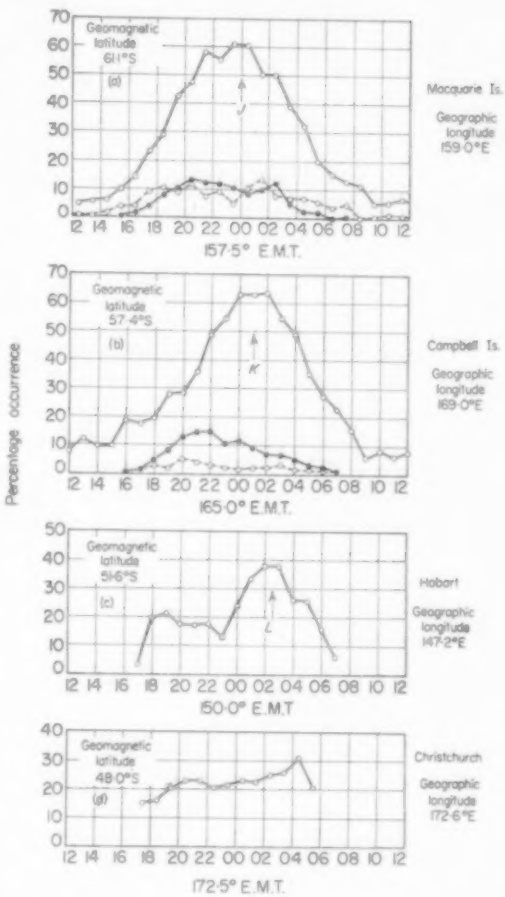


Fig. 1. Diurnal curves of percentage hourly occurrence of E_s. (i) $f + l$ type E_s : O——O. (ii) a type E_s : ●——●. (iii) r type Es: O— — —O

Figures 1(a) and 1(b) also show the diurnal variations of *a* type E_s and *r* type E_s. The significance of these two types has been discussed[2]. A very small percentage occurrence of *a* type E_s was found at *H* for some hours, and has been omitted. *a* type E_s is absent from the *CHCH* hourly data. *r* type E_s was not present in the hourly data for either of these stations. During sunlit hours the presence of *l* type E_s has been included for all four stations.

3. DISCUSSION OF E_s DIURNAL OCCURRENCE

An examination of Fig. 1 shows the presence of one major night-time peak in the percentage occurrence of *f* type E_s at three of the four stations. The local time of occurrence of this peak advances as the geomagnetic latitude becomes smaller. This behaviour was observed for the Auroral Belt stations by Penndorf and Coroniti. The presence of this peak is detectable at *H* (geomagnetic latitude 51·6 °S), in the equinoctial months. This indicates that the agency responsible for the occurrence of Auroral Belt type E_s extends well into latitudes below the visual auroral belt. Its presence at *CHCH*, however, is not detected by the present analysis. The minimum width of the zone over which this agency produces night-time sporadic *E* can be placed between geomagnetic latitudes 50–70 degrees.

The maximum percentage occurrence of approximately 60 per cent for all seasons at *MI* and *CI*, bearing in mind the restrictions on the data which have been explained above, shows that the agency which causes this type of E_s must be present to some degree on almost all days. Its presence cannot be explained solely by the processes responsible for visual auroral displays and the accompanying magnetic disturbance.

The diurnal occurrence of the peak near midnight indicates some form of solar control, with an intermediate process which favours the midnight meridian.

The dependence of the temporal occurrence of the peak on geomagnetic latitude at each observing site suggests magnetic control which is dependent on the curvature and intensity of the earth's magnetic field.

4. THEORY OF THE OCCURRENCE OF AURORAL BELT TYPE E_s

The following assumptions are made to describe the process which produces the observed behaviour of Auroral Belt type E_s.

(a) The night-time occurrence of this type of E_s results from the ionisation caused by the influx of predominantly negatively charged particles, in this case electrons, into the *E*-region of the ionosphere.

(b) The electrons are derived from a region of trapped particles occupying the space enclosed by the surfaces generated by magnetic field lines which meet the *E*-region of the ionosphere at geomagnetic latitudes of approximately 50 and 70 degrees.

(c) The geomagnetic field is distorted in such a way that electrons in the trapping region, which receive additional energy from disturbances on the day side of the earth, travel round the region until they reach the point where the magnetic flux will no longer reflect them. They are then precipitated into the *E*-region.

Measurements carried out during rocket flights and reported by Van Allen *et al.*[5] show the probable existence of such a region of trapped electrons. Since the electrons in this region will have a component of motion eastwards, they continually travel around the earth. At any point in its motion along a given line of force, an electron will move under an eastward component of force which is a function of the curvature of the field line and the magnetic intensity at that point. In the equilibrium state of the region there will be an upper energy limit for the particles contained, a limit determined by the minimum magnetic flux existing at some level in the ionosphere. Particles which have energies greater than this upper limit will be contained in the region for only part of their circuit of the earth before being lost to the *E*-region.

It has been assumed in (c) that the region of minimum magnetic flux is situated near the midnight magnetic meridian and that its position and form is under solar control.

VOL 2 1959

Figure 2 gives a diagrammatic section of the earth's field and the trapping region for a centred dipole field at "geomagnetic noon" during the southern summer solstice. A and A', respectively, denote the northern and southern lower borders of the trapping region. It has been assumed in (b) that the electrons on the

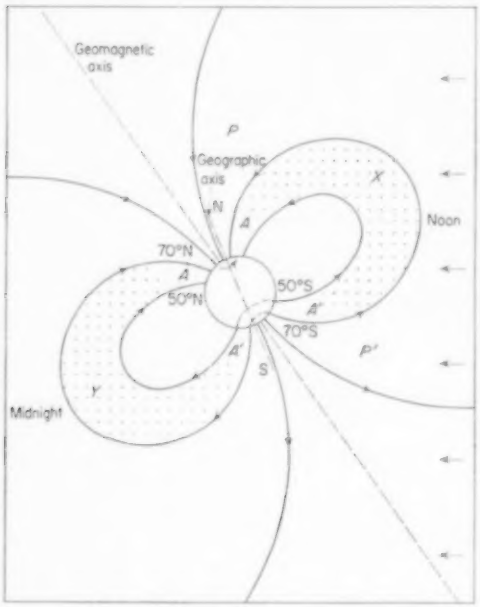


Fig. 2. Diagrammatic section of the geomagnetic field for a centred dipole at "geomagnetic noon" in the southern summer solstice, showing the region of trapped particles, A-A, and the zone of occurrence of the night-time peak in f type E_s.

day side may receive additional energy from disturbances arriving near the region X. The electrons will then move eastwards around the field till they reach the region Y, on the midnight side, when they are precipitated at A and A'. The eastward deflection of the electrons on the final trajectory displaces them, so that on arrival at the E-region they appear to be displaced in time. This eastward displacement increases as the curvature of the lines of forces and the magnetic field strength increase on the low latitude side of the trapping region.

5. THEORY OF THE OCCURRENCE OF THULE TYPE E_s

Penndorf and Coroniti have shown that Thule type E_s has its maximum occurrence near "geomagnetic noon" and is dependent on universal time. Further it has its maximum occurrence in summer.

In Fig. 2 it is seen that these facts are most readily satisfied if it is assumed that electrons approaching the earth from the direction of the sun are guided along the field lines in the region P' and precipitate in the E-region about the "geomagnetic noon" meridian at geomagnetic latitudes greater than 70 degrees.

As in the previous case, the electrons receive an easterly displacement so that they appear to an observer on the earth to be displaced in time. The curvature of the magnetic lines of force becomes less as the magnetic pole is approached, and the mean easterly displacement will therefore be less than for electrons which produce Auroral Belt type E_s.

CONCLUSION

On the basis of the theory discussed in this paper, measurements of the occurrence of E_s in the polar regions can provide further information about the energies of particles which are guided into the E-region by the geomagnetic field. These measurements can also provide information on the form and extent of the geomagnetic field and its behaviour when it interacts with solar particles.

Acknowledgments—Some of the ionospheric observations, used in the analysis, have been carried out using facilities provided by the Australian National Antarctic Research Expedition at Macquarie Island.

REFERENCES

1. R. Penndorf and S. C. Coroniti, *J. Geophys. Res.* **63**, 789 (1958).
2. J. E. Shaw, *Planet. Space Sci.*, to be published.
3. J. A. Thomas and E. K. Smith, *J. Atmos. Terr. Phys.* **13**, 314 (1959).
4, 6. *Ionospheric Prediction Service*, Commonw. of Aust., Dept. Interior, Ser. D (1957-58).
5, 7. *Christchurch Geophysical Observatory*, New Zealand, D.S.I.R., Form 101 (1957-58).
8. J. A. van Allen, G. H. Ludwig, E. C. Ray and C. E. McIlwain, *State Univ. Iowa Res. Rep.* (May 1958).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 60–70. Printed in Great Britain.

EXCITATION OF O_2 ATMOSPHERIC BANDS IN THE AURORA*

L. WALLACE and J. W. CHAMBERLAIN

Yerkes Observatory, University of Chicago, Williams Bay, Wisconsin

(*Received 5 September* 1959)

Abstract—Spectra of the aurora show a wide disparity between the rotational temperatures of the two observed Atmospheric bands, $b\,^1\Sigma_g^+ \rightarrow X\,^3\Sigma_g^-$, with $T_{rot}\,(0\text{–}1) \approx 200°K$ and $T_{rot}\,(1\text{–}1) \gtrsim 700°K$. As the rotational structure of the bands is not resolved, it is not absolutely certain that the temperature anomaly is real, but known features in the auroral spectrum would not produce the effect through blending. Fluorescence in the 0-0 band provides a possible but improbable explanation. Another alternative to explain the data is excitation of O_2 by energy transfer from $O\,(^1D)$ (upper term of the red lines) followed by vibrational deactivation. In the latter case the relative intensities of the O_2 bands and the [O I] red lines gives an indication of the rate coefficients for the relevant reactions, which are then used to predict the deactivation probability for $O\,(^1D)$ as a function of height. The reactions discussed give a satisfactory explanation of the absence in the airglow of atmospheric O_2 bands from vibrational levels $v' \geq 1$, and predict virtually complete suppression of the [O I] red lines at 100 km in the airglow. The airglow O_2 emission gives $T_{rot}\,(0\text{–}1) = 183°K$.

1. INTRODUCTION

The atmospheric system of O_2, $b\,^1\Sigma_g^+ \longrightarrow X\,^3\Sigma_g^-$ was first discovered in the airglow and later in the aurora by Meinel[1,2]. He found the 0–1 band at 8645 Å in emission, but the 0–0 band, which should be much stronger, is reabsorbed (and possibly converted into 0–1, 0–2, etc.) by the great amount of O_2 below the emitting region. Meinel reported an airglow rotational temperature of $T_{rot}\,(0\text{–}1) = 150°K$ (see, however, a correction reported in Section 2).

In the aurora the presence of the 1–1 atmospheric band at 7708 Å was first suggested by Kvifte[3] and later confirmed by Chamberlain *et al.*[4], who estimated that this band had about one-half the intensity of 0–1. They also estimated $T_{rot}\,(0\text{–}1) \approx 200°K$ in the aurora, but because the 1–1 band is partially blended with other features, made no measurement of $T_{rot}\,(1\text{–}1)$.

From a number of infrared auroral spectra of moderate dispersion, we have been able to make some improved measurements of the relative intensities of these bands and their rotational temperatures, which we report in Section 2. While the relative intensities seem consistent with excitation of these bands by electron impact, a disparity in the rotational temperatures of the two bands cannot be understood on this basis. Also it does not appear that fluorescence in the 0–0 band is responsible for the temperature difference.

It does not seem likely that the high temperature we derive is produced by the N_2 first positive band that overlaps the O_2 band or by other overlying emissions. Still, the rotational structure is not resolved and one must bear in mind the possibility that unknown lines may be blended with the band and distort the profile. Nevertheless, this seems an artificial way to explain the data and we have inquired whether the effect may be a natural consequence of the excitation mechanism.

We propose in Section 3 that the excitation of these bands occurs predominantly by energy

* The research reported in this paper was supported in part by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF19 (604)-3044 with the University of Chicago.

transfer from the upper term of the red lines, $O(^1D)$, followed by vibrational deactivation. These mechanisms allow an explanation of the difference in rotational temperatures, and the relative intensities of the bands provide crude information on the relevant rate coefficients.

Since the mechanism populating O_2 also deactivates $O(^1D)$, we estimate in Section 4 the deactivation probability for this term. These results predict virtually complete suppression of λ 6300 in the 100-km region of the airglow, even with the large amount of excitation expected for this line. We shall show that the rather large transition probability for $b\,^1\Sigma_g^+ \longrightarrow a^1\Delta_g$ hypothesized by Seaton[5] is neither necessary to explain the deactivation of $O(^1D)$ nor justified by the observations.

2. OBSERVATIONS

Five auroral and four airglow spectra, taken with the Meinel $f/0{\cdot}8$, 9-in. aperture spectrograph (described by Chamberlain[6]), were chosen for analysis. In some cases a single plate was used as an example of both an auroral and an uncontaminated airglow spectrum. This was possible because these plates were taken with the northern sky imaged on one part of the slit and the eastern sky imaged on the other part. In several instances a single plate showed a faint auroral spectrum on the northern section and an apparently uncontaminated airglow spectrum on the eastern section. For the purpose of this work, the presence of λ 7774 of O I, one of the strongest lines in the near infrared auroral spectrum, was used to indicate an aurora. Intensity calibrations, consisting of spectra of a standard incandescent lamp taken with a step filter across the entrance slit of the spectrograph, were obtained at the time the sky spectra were photographed. The spectra were reduced in the usual manner and corrected for the monochromatic sensitivity of the spectrograph and the emulsion (Eastman *I–N*).

We constructed model profiles, with the wavelengths of the rotational lines of the 0–1 and 1–1 bands calculated on the basis of the analysis of this system by Babcock and Herzberg[7]. The ^RR and ^PP branch wavelengths were calculated from rotational constants and the ^RQ and ^PQ wavelengths were determined to sufficient accuracy for this work from the approximate relation

$$^PQ(K) - {}^PP(K) = {}^RQ(K) - {}^RR(K) = 2{\cdot}0\ \mathrm{cm}^{-1}.$$

The intensities of the rotational lines were obtained from the Boltzmann distribution of populations (see Herzberg[8]) and the theoretical line strengths (or intensity factors) computed by Schlapp[9]. The latter are

$$i(^PP) = \frac{1}{2}(J+2);\ i(^PQ) = \frac{1}{2}\left(J+\frac{3}{4}\right);\ i(^RR) = \frac{1}{2}(J-1);\ i(^RQ) = \frac{1}{2}\left(J+\frac{1}{4}\right);$$

where the J's refer to the upper state. These strengths for the ^PQ and ^RQ branches are approximations differing from the exact theoretical expressions by no more than one per cent. The intensities of the individual rotational lines were distributed according to the observed instrumental profile and then summed to obtain the model profiles given in Figs. 1, 2 and 3. The rotational temperatures refer to equilibrium in the upper state.

In order to determine a rotational temperature from the airglow 0–1 band observations, both the models and the observations were first normalized to the total band intensity, as illustrated in Fig. 1. Then the sum of the squares of the differences between the normalized observed intensities and model intensities taken at 4 Å intervals were calculated for each observed profile to determine which temperature model gave the smallest sum. The average of these temperatures for the smallest sums is $183°\mathrm{K} \pm 7°$ (m.e.) where the mean error was obtained from the scatter in the temperatures. The uncertainty in the temperature determination for a single observed profile is estimated to be of the order of 10° or 20°. Consequently there is no evidence in the present observations for a real temperature variation.

Meinel[1] obtained $T(0-1) = 150°\mathrm{K} \pm 20$ for the airglow. However, when Meinel's observations are compared with the present models,

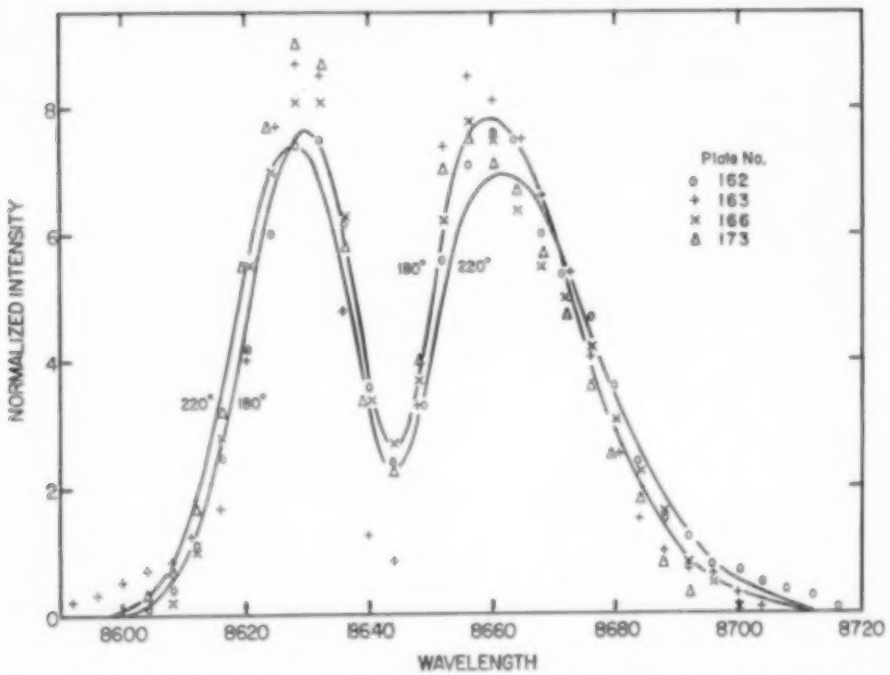


Fig. 1. Observed and model intensity profiles of the 0–1 band in the airglow. The solid lines indicate the model profiles. The plate numbers are identified on Table 1.

they indicate $T(0-1)=160°K$. This 10° difference is due mainly to Meinel's misplacement of the ^rQ branch by 4 cm^{-1} but we have also used slightly different strengths. Consequently, the present observation of $T(0-1)=183°K\pm7°$ is to be compared with that of $160°K\pm20°$ determined from Meinel's observations. Dufay and Dufay's estimate[13] of 130° from low-dispersion spectra is probably not sufficiently accurate to indicate a real temperature difference.

The 0–1 band auroral profiles were treated in the same way as the airglow profiles except that the comparison between the observed and model profiles was limited to the wavelength range 8606–8666 Å because of overlap with other spectral features. These include multiplet* 1 of NI (8680–8748 Å), the 2–1 band of the First Positive system of N_2 in the region 8670–8722 Å (see Chamberlain and Meinel[12]), the P_1 (7) and and P_2 (7) lines of the 6–2 band of OH (λλ 8597 and 8588; see Herman and Hornbeck[13]), and occasionally λ 8629 of multiplet 8 of NI. Higher rotational members of the 6–2 band of OH should be faint compared with the O_2 band. If we exclude plate No. 131, which shows λ 8629 of NI, the remaining five auroral plates listed in Table 1 showed the region 8606–8666 Å apparently free of overlap. Using this region for the least squares fit, we found the average temperature for the 0–1 auroral bands to be $192°K\pm6°$.

To test the possibility that λ 8629 of NI was affecting the R branch, the observations were reduced again with the region 8624–8636 Å omitted from the least-squares determination of the temperature. The average of these reductions gave $T(0-1)=197°K\pm8°$. With the inclusion of the profile from plate No. 131 this average was altered to $195°K\pm8°$. From the

* The multiplet numbering convention is that adopted by Moore[11].

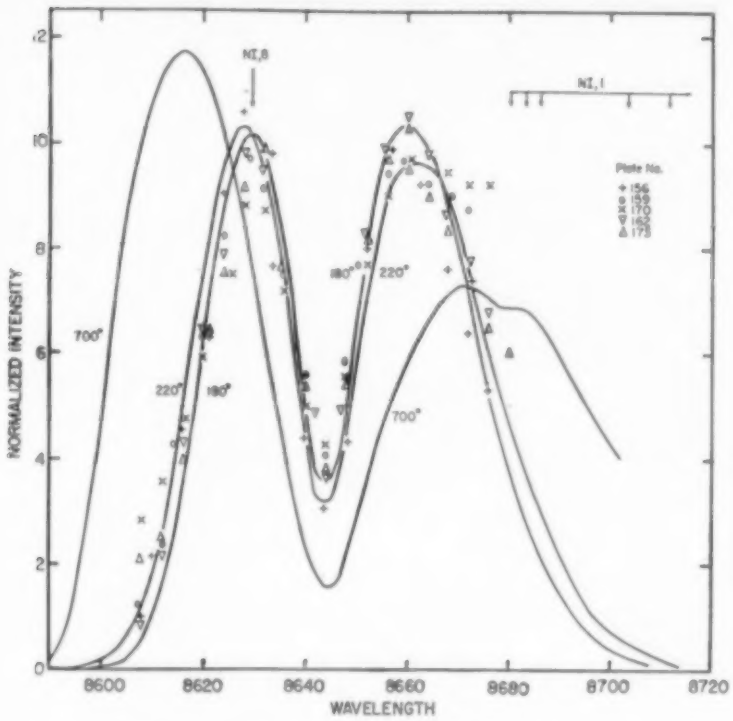


Fig. 2. Observed and model intensity profiles of the 0–1 band in the aurora.

consistency of these determinations we conclude that 195°K ± 10° is a good measure of the rotational temperature of the upper state of the 0–1 band. This is very nearly the same as the 200° value estimated by Chamberlain *et al.*[4] for the aurora and only slightly higher than the value of 183°K ± 7° obtained above for the airglow.

The observations of the 1–1 band cannot be converted into temperatures in the manner of the 0–1 band because only the region of the *R*-branch maximum appears to be free of overlap with the 2–0 First Positive band of N_2 (in the range 7690–7753 Å). The observed profile showing the least overlap with the 2–0 First Positive band is shown in Fig. 3. For this one observation a temperature of 600°K appears to be valid, provided that the 2–0 First Positive band does not influence the profile shortward of 7695 Å. The other 1–1 band observations show greater overlap with the 2–0 First Positive band and only the intensity maxima seem free of overlap. The position of this maximum was found to be 7686·3 ± 0·6 Å (m.e.). This measure, indicated on Fig. 3, requires a temperature of 1000° to 2000°K. It is clearly difficult to obtain an accurate temperature for this band. The best estimate seems to be $T_{rot}(1\text{–}1) \geqslant 700°$. Regardless of the difficulty in assigning a temperature to the 1–1 band, it is clear from Figs. 2 and 3 that T_{rot} (1–1) is considerably greater than T_{rot} (0–1).

The relative intensities of the 1–1 and 0–1 bands can be used to estimate the relative populations integrated along the line of sight, of the first two vibrational levels of the upper state, i.e. $\mathcal{N}(v'=0)/\mathcal{N}(v'=1)$. We shall use $\mathcal{I}$ to denote an integrated intensity over an entire line

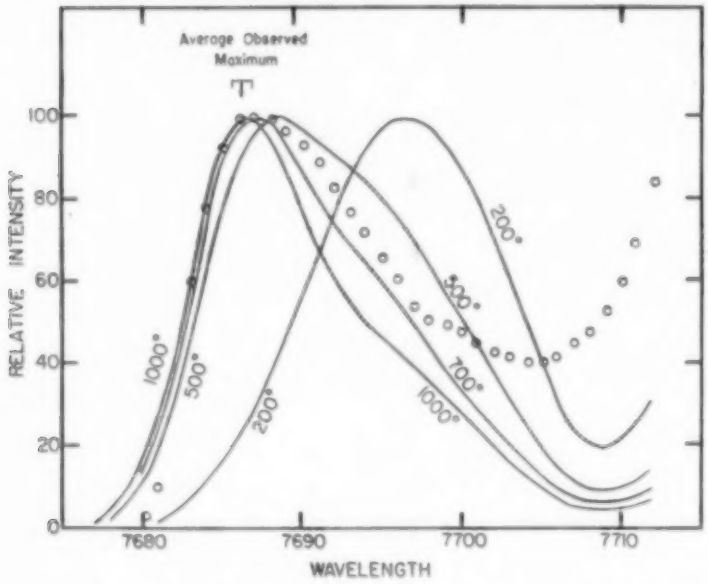


Fig. 3. Observed and model intensity profiles of the 1–1 band in the aurora. The observed profiles are seriously blended with other features and only appear to be free of overlap near the intensity maximum of the *R*-branch. The observed profile least affected by overlap is indicated by open circles. For the other profiles only the mean wavelength of the intensity maxima and the mean deviation of that wavelength are indicated.

or band profile. The intensity ratios $\mathcal{I}(0\text{–}1)/\mathcal{I}(1\text{–}1)$, $\mathcal{I}(1\text{–}1)/\mathcal{I}(\lambda\,7774)$ and estimates of $\mathcal{I}(2\text{–}2)/\mathcal{I}(1\text{–}1)$, along with $\mathcal{I}(\lambda\,7774)/\mathcal{I}(\mathrm{OH})$ which serves as a rough estimate of the brightness of the aurora, and $\mathcal{I}(0\text{–}1)/\mathcal{I}(\mathrm{OH})$ which indicates the enhancement of the 0–1 band, are given in Table 1. We took $\mathcal{I}(\mathrm{OH})$ as the intensity of the $P_1(2)$ line of the 6–2 band of OH, which on our spectra appears to be free of overlap with auroral features.

The integrated intensities of the O_2 bands were corrected for overlapping features by extrapolation based on the 195°K model for 0–1 band and the 700°K model for the 1–1 band. If the correct 1–1 band rotational temperature is 600°, $\mathcal{I}(1\text{–}1)$ given in Table 1 will be 7 per cent too low, and if the correct temperature is 1000°, it will be 9 per cent too high. The errors introduced into $\mathcal{I}(0\text{–}1)$ by possible errors in the selection of $T_{\mathrm{rot}}(0\text{–}1)$ will be small compared with those in $\mathcal{I}(1\text{–}1)$ and both are less than the estimated 20 per cent photometric error.

The ratios $\mathcal{I}(0\text{–}1)/\mathcal{I}(1\text{–}1)$ were used with the calculated Franck–Condon factors given by Fraser *et al.*[14] to obtain the relative populations. Since there are no measurements of the relative intensities of bands originating from a common upper level in this system, we have assumed that the electronic transition moment, $R_e(r)$, is constant over the range or r valid for the 0–1 and 1–1 bands. The possible error this assumption might introduce is probably small, since in other band systems studied, $R_e(r)$ has been found to be slowly varying (Nicholls[15]). However, the difference between Kvifte's[3] observed estimate (with a laboratory discharge source) that $\mathcal{I}(0\text{–}1)/\mathcal{I}(0\text{–}0)\approx 0{\cdot}02$ and the value of 0·04 obtained from the Franck–Condon

factors of Fraser *et al.* with $R_e(r)$ assumed constant might indicate a factor of 2 variation in R_e^2. If so, the values of $\mathcal{N}(v'=0)/\mathcal{N}(v'=1)$ in Table 1 are a factor of 2 too small. In any event it is clear that the population ratio is decreased by more than a factor of 8 in the aurora compared with the airglow.

3. O_2 EXCITATION AND DEACTIVATION MECHANISMS

The auroral observations might be interpreted in terms of excitation occurring through energy transfer by (Bates and Dalgarno[16], Seaton[17])

$$O(^1D)+O_2(X^3\Sigma_g^-, v''=0)\longrightarrow O(^3P)+O_2(b^1\Sigma_g^+, v'\leqslant 2). \quad (1a)$$

This reaction is in near resonance for $v'=2$, but is still energetically possible for lower values of v'. Hence deactivation of $v'=2$ may occur by the inverse reaction,

$$O_2(b^1\Sigma_g^+, v'=2)+O(^3P)\longrightarrow O_2(X^3\Sigma_g^-, v''=0)+O(^1D), \quad (1b)$$

but for $v'=0$ and 1, the latter reaction is not energetically possible without an exceptionally large amount of kinetic energy. All values of $v'>0$ might be deactivated, however, by a transfer of electronic energy only (Bates[18]):

$$O_2(b^1\Sigma_g^+, v'>O)+O_2(X^3\Sigma_g^-, v''=0)\longrightarrow O_2(X^3\Sigma_g^-, v''>O)+O_2(b^1\Sigma_g^+, v'=0) \quad (2)$$

Either electronic or vibrational deactivation of O_2 might occur by atom–atom interchange (Bates[19]):

$$O_2(b^1\Sigma_g^+, v')+O(^3P)\longrightarrow O(^3P)+O_2(a) \quad (3)$$

with $a=b^1\Sigma_g^+$ $v<v'$ or $a=$ a lower electronic state. While this type of process is likely responsible for suppressing the Herzberg O_2 bands in the airglow, the activation energy required to break the bond for $b\ ^1\Sigma_g^+$ probably inhibits the process from affecting the Atmospheric system. In any event, we shall neglect it in our present discussion.

The main alternative to excitation of O_2 by reaction (1*a*) is probably direct electron impact, which is presumably the dominant mechanism exciting $O(^1D)$. However, this mechanism does not seem consistent with the disparity in rotational temperatures of the 0–1 and 1–1 bands. Electron excitation should produce little or no change in the rotation of a molecule (Massey

Table 1. Relative Photon Intensities and Relative Populations of Vibrational Levels†

Plate	Date	$\mathcal{I}(\lambda 7774)/\mathcal{I}(OH)$	$\mathcal{I}(0\text{–}1)/\mathcal{I}(OH)$	$\mathcal{I}(1\text{–}1)/\mathcal{I}(\lambda 7774)$	$\mathcal{I}(2\text{–}2)/\mathcal{I}(1\text{–}1)$	$\mathcal{I}(0\text{–}1)/\mathcal{I}(1\text{–}1)$	$\mathcal{N}(v'=0)/\mathcal{N}(v'=1)$	$T(0\text{–}1)$ (°K)
159	5 Dec. 1958	4·7	9·0	1·1	<0·18	1·6	28·0	190
170	10 Jan. 1959	4·5	19·0	2·4	<0·33	1·7	30·0	200
131	31 Mar. 1958	3·4	19·0	2·3	<0·16	2·4	43·0	185
162	9 Dec. 1958	1·5	6·0	1·7	<0·38	2·4	43·0	185
173	16-17 Jan. 1959	0·8	2·5	2·3	<0·19	1·3	23·0	190
156	18 Nov. 1958	0·4	4·2	1·3	<0·9	9·0	160·0	195
162	9 Dec. 1958	—	4·2	—	—	> 6·0	>100·0	190
163	10 Dec. 1958	—	3·5	—	—	> 7·0	>110·0	175
166	13 Dec. 1958	—	4·4	—	—	>12·0	>200·0	185
173	16-17 Jan. 1959	—	1·6	—	—	> 3·3	> 60·0	180

† The first six entries in this table are examples of auroral spectra. The remaining four, which showed no auroral features, are given as examples of airglow spectra. $\mathcal{I}(OH)$ is the intensity of the P_1 (2) line of the 6-2 band of OH.

and Burhop[20], p 276), at least when the molecule is not simultaneously excited and ionized by the collision. Hence both bands should show T_{rot} characteristic of the gas temperature. Also it seems reasonable to suppose that reaction (1*a*) would dominate electron impact from the considerations that the electronic cross section for exciting O_2 is not likely to be much larger than that for exciting O; that in most of the emitting region the O abundance probably will exceed that of O_2; and that at the lower auroral altitudes most of the $O(^1D)$ excitation seems to be deactivated, probably by reaction (1*a*) (Seaton[5]).

One possible explanation for the disparity in rotational temperatures is that the 0–1 band might arise predominantly from fluorescence, following absorption of the more intense 0–0 band (Bates[18]). The 1–1 band would receive very little enhancement resulting from absorption in the 1–0 band, because of the relative transition probabilities involved. Therefore, T_{rot} (0–1) might pertain predominantly to the temperature in the region where fluorescence occurs, whereas T_{rot} (1–1) would apply to the rotational distribution (not necessarily to the kinetic temperature) in the region of the auroral excitation.

The fluorescence mechanism occurs in the general region where the vertical optical thickness in the individual rotational lines becomes the order of unity, which occurs at heights around 40 to 50 km for the strongest O_2 lines. The radiative-transfer problem has been treated by Chamberlain[21]. The temperature in this region exceeds 250°K; hence the low rotational temperature in the 0–1 airglow band argues against the fluorescence mechanism. It appears that at these low altitudes excited O_2 molecules are deactivated to the ground electronic state and that fluorescence cannot therefore explain the auroral temperatures. (It may be noted in passing that were fluorescence important, an aurora seen near the horizon would show the 0–1 band at a much lower altitude than the 1–1 band and other auroral emissions.)

Consider now the excitation produced by the above reactions. It is not known how important is resonance in reaction (1*a*) but let us consider the reaction rate for excitation of a level v' to be $\alpha_{1a}(v')$. We shall take the rate coefficient α_2 to be independent of v' and write $O(^1D) \equiv O^*$ and $O_2(^1\Sigma_g^+, v') \equiv O_2^*(v')$; the number density of particles will be written with brackets, as in $[O_2^*(v')]$. Qualitatively we might expect $[O_2^*(0)]$ to be greater than $[O_2^*(v'>0)]$ because of reaction (2) alone. But for $v'=2$ there is an additional deactivation by reaction (1*b*).

If the excitation of $v'=1$ is produced by reaction (1*a*) it may well have T_{rot} much greater than the gas temperature; there is no reason to expect it to approximate the gas temperature. But $v'=0$ is evidently populated largely by reaction (2), where initially the O_2^* molecule has $v'=1$. (Contributions from $v'=2$ will be negligible because, as noted above, $[O_2^*(2)] \ll [O_2^*(1)]$.) Since the reaction involves only electronic transfer, the final state $O_2^*(0)$ has both the vibrational ($v'=0$) and rotational ($T_{rot} \approx T_{kin}$) structures of the initial ground-state molecule, $O_2(X^3\Sigma_g^-)$. Also, molecules that are excited directly to $v'=0$ may still have their rotational temperature lowered by collisions similar to reaction (2), wherein electronic transfer occurs. As long as ordinary thermal collisions (which redistribute the rotational populations) do not occur before an $O_2^*(0)$ molecule radiates, the rotational intensities should be given by the kinetic temperature and the rotational constant (B''_0) for the ground state. Thermal collisions will change the distribution. But in practice $B'_0 \approx B''_0$ so it matters little which constant is used. For the 1–1 band the rotational temperature will not have a direct bearing on the kinetic temperature, so long as this band is emitted only by molecules that have not suffered many collisions, and especially no collisions of the type (2).

We consider these matters now in a more quantitative way. If $O_2^*(2)$ is populated only by reaction (1*a*), and depopulated by (1*b*), (2) and radiation with a transition probability A_{bx}, its poulation in a steady state is

$$\frac{[O_2^*(2)]}{[O_2]} = \frac{[O^*]\,\alpha_{1a}(2)}{[O]\,\alpha_{1b} + [O_2]\,\alpha_2 + A_{bx}}. \qquad (4)$$

For $v'=1$ deactivation by (1*b*) cannot occur so we have

$$\frac{[O_2^*(1)]}{[O_2]} = \frac{[O^*]\,\alpha_{1a}(1)}{[O_2]\,\alpha_2 + A_{bX}}. \qquad (5)$$

Finally, for $v'=0$ reaction (2) can add to the excitation giving

$$\frac{[O_2^*(0)]}{[O_2]} = \frac{[O^*]\,\alpha_{1a}(0) + [O_2^*(1)]\,\alpha_2}{A_{bX}}, \qquad (6)$$

where we neglect the term $[O_2^*(2)]\,\alpha_2$ in the numerator because of the observed fact that $[O_2^*(2)] \ll [O_2^*(1)]$.

These equations yield the relative populations of the upper levels of the observed emissions. Combining equations (5) and (6) gives

$$\frac{[O_2^*(0)]}{[O_2^*(1)]} = \frac{[O_2]\,\alpha_2}{A_{bX}}\left\{\frac{\alpha_{1a}(0)}{\alpha_{1a}(1)} + 1\right\} + \frac{\alpha_{1a}(0)}{\alpha_{1a}(1)}. \qquad (7)$$

Equations (4) and (5) provide the ratio

$$\frac{[O_2^*(2)]}{[O_2^*(1)]} = \frac{\alpha_{1a}(2)}{\alpha_{1a}(1)}\,\frac{\{[O_2]\,\alpha_2 + A_{bX}\}}{\{[0]\,\alpha_{1b} + [O_2]\,\alpha_2 + A_{bX}\}}, \qquad (8)$$

and equation (5) may be rewritten as

$$\frac{[O_2^*(1)]}{[O^*]} = \frac{\alpha_{1a}(1)\,[O_2]}{\alpha_2\,[O_2] + A_{bX}}. \qquad (9)$$

By comparing these equations with the observed intensities we can obtain some rough estimates of the rate coefficients that are required with our interpretation. We note first of all that in the airglow the Atmospheric system is probably excited mainly by three-body association, where the rate coefficient for formation of a molecule in the state $b\,^1\Sigma_g^+$ and level v' is $\beta(v')$. The absence of bands with $v'>0$ in the airglow, even though the higher levels should be excited, indicates that deactivation by (2) is more rapid than spontaneous emission, or $[O_2]\,\alpha_2 > A_{bX}$. Then statistical equilibrium considerations of the sort used to derive the above equations give for the airglow

$$\frac{[O_2^*(0)]}{[O_2^*(1)]} = \frac{[O_2]\,\alpha_2}{A_{bX}}\,\frac{\sum_{v'=0}\beta(v')}{\beta(1)}, \qquad (10)$$

where the summation is carried over all levels.

The data in Table 1 indicate that the ratio expressed by equation (10) exceeds 10^2. Although the relative rate coefficients $\beta(v')$ are not known, they probably do not change rapidly with increasing v', and $\Sigma\beta(v')/\beta(1)$ may be of the order of 10. With $A_{bX}=0{\cdot}12$ sec^{-1} and $[O_2]\sim 10^{12}$ cm^{-3} (with Nicolet's[22] model the maximum emission rate would be in the 95 km region, as observed) the airglow observations require that $\alpha_2 > 10^{-12}$ cm^3/sec, which does not seem to be a severe condition. (A reaction proceeding with the gas-kinetic cross section would have a coefficient of about 3×10^{-10} cm^3/sec.)

Since the 1–1 band is characterized by a temperature well in excess of the kinetic, thermal collisions that redistribute the rotational levels must occur not more than a few times during the lifetime of $v'=1$. As this lifetime is governed by reaction (2) at the lower auroral altitudes, we find that α_2 must be the order of the gas-kinetic rate. This condition is more serious; the validity if our entire interpretation requires, however, that reaction (2) proceed very efficiently. No experimental data on such a reaction is available.

Thus we shall assume that $[O_2]\,\alpha_2 > A_{bX}$ for a considerable range above the airglow region—at least up to 150 km. Therefore we see from equation (9) that the intensity ratio of the 1–1 band relative to [O] red lines should be nearly constant over the lower auroral region, which is in reasonable agreement with Omholt's[23] observations. Table 1 gives $\mathcal{I}(1\text{–}1)/\mathcal{I}(7774) \approx 1{\cdot}8$. Combining this with Omholt's[23] measurement for the lower part of auroral displays, $\mathcal{I}(7774)/\mathcal{I}(6300\text{–}64) = 0{\cdot}27$, and using for the transition probability of the red lines $A_{DP} = 0{\cdot}91\times 10^{-2}$ sec^{-1}, we obtain the mean ratio $[O_2^*(1)]/[O^*] \approx \alpha_{1a}(1)/\alpha_2 \approx 0{\cdot}04$.

The absence of the 2–2 band in the aurora must be explained in terms of equation (8). Seaton[5] has pointed out that detailed-balancing considerations require that $\alpha_{1a}(2) = 0{\cdot}6\,\alpha_{1b}$. Hence, we cannot explain the small ratio of equation (8) by assuming that $[O]\,\alpha_{1b} \gg [O_2]\alpha_2$, for we should then find from equation (8) that in the low auroral region $\alpha_2 < \alpha_{1a}(1)$, which contradicts the result in the preceding paragraph.

Therefore, we must assume that $[O_2]\alpha_2 > [O]\alpha_{1b}$ in the lower auroral region. In that case equation (8) gives $[O_2^*(2)]/[O_2^*(1)] \approx \alpha_{1a}(2)/\alpha_{1a}(1) < 0{\cdot}2$. Thus we conclude that resonance does not seem to be so important in reaction (1*a*) after all, and that reaction (1*b*) is less important in deactivating $v' = 2$ than is reaction (2), at least at the lower auroral altitudes where O_2 is relatively abundant.

An estimate of the ratio $\alpha_{1a}(0)/\alpha_{1a}(1)$ from equation (7) is not so straightforward, since the intensity ratio is strongly dependent on the altitude of excitation. The strong concentration of the 0–1 emission relative to 1–1, toward the lower border of the aurora as indicated by equation (7), is probably the explanation of the low auroral value of $T_{rot}(0\text{–}1)$. Relatively little of the total red-line excitation would have to occur at 100 ± 5 km for that region to dominate in producing the 0–1 emission.

The abundance ratio $[O_2^*(0)]/[O_2^*(1)]$ could be smaller in the aurora than in the airglow either because $v' = 1$ is more favored in the auroral excitation mechanism or (more likely) as a height effect, with $v' \geq 1$, not so readily deactivated in the aurora as in the airglow. In the aurora sometimes we find (see Table 1) the abundance ratio $[O_2^*(0)]/[O_2^*(1)] < 30$. Hence by equation (7) this is an upper limit on the ratio of the rate coefficients, $\alpha_{1a}(0)/\alpha_{1a}(1)$, as well. And since all our auroral spectra show $T_{rot}(0\text{–}1) \approx 200°\text{K}$, most of the 0–1 excitation evidently occurs at low enough altitudes that vibrational deactivation (2) contributes a fair portion of the $[O_2^*(0)]$ population. Thus it appears that $\alpha_{1a}(0)/\alpha_{1a}(1) \ll 30$.

Table 2. Summary of Transition Probabilities and Rate Coefficients

$A_{b\Sigma}$	$= 0{\cdot}12\ \text{sec}^{-1}$
A_{DP}	$= 0{\cdot}91 \times 10^{-2}\ \text{sec}^{-1}$
α_2	$\sim 10^{-10}\ \text{cm}^3/\text{sec}$
$\alpha_{1a}(1)$	$\sim 4 \times 10^{-12}\ \text{cm}^3/\text{sec}$
$\alpha_{1a}(2)$	$< 0{\cdot}8 \times 10^{-12}\ \text{cm}^3/\text{sec}$
α_{1b}	$< 1{\cdot}3 \times 10^{-12}\ \text{cm}^3/\text{sec}$
$\alpha_{1a}(0)$	$\ll 1{\cdot}2 \times 10^{-10}\ \text{cm}^3/\text{sec}$

Table 2 summarizes the conditions on the rate coefficients that seem necessary for these reactions to explain the data.

4. DEACTIVATION OF [O I] RED LINES

Reaction (1*a*) is also of interest in explaining the absence of λ 6300 airglow emission from O (^1D) around 100 km, where the green line is formed, and deactivation of O (^1D) in the aurora at low altitudes. Seaton[5] noted that reaction (1*b*) would re-excite O (^1D) unless there were a more efficient means than radiation by the $b \longrightarrow X$ transition to remove $O_2^*(2)$ molecules. Therefore he postulated that the transition $b\,^1\Sigma_g^+ \longrightarrow a\,^1\Delta_g$ could occur with a transition probability $A_{ba} \sim 1\ \text{sec}^{-1}$.

If resonance were indeed important in (1*a*), the deactivation of $O_2^*(2)$ by reaction (2) would serve the same purpose, however. And if resonance is not important, as the ratio of the 2–2 and 1–1 band intensities suggests, there is no problem in re-exciting O (^1D). These processes, therefore, seem quite capable of explaining red-line deactivation at low altitudes, and, indeed, the use of Bates' mechanism, reaction (2), at the 100 km region seems necessary to explain the absence of the higher vibrational levels of O_2 ($b\,^1\Sigma_g^+$) in the airglow. Reaction (2) becomes less important with height and were O_2 excited at higher altitudes in the aurora, one would expect to see the emission from $v' = 1$ become relatively stronger.

Seaton's postulate that $A_{ba} \gg A_{b\Sigma}$ not only seems unnecessary but is also probably unjustified on the basis of the intensity that should appear in the nightglow Infrared Atmospheric system, $a\,^1\Delta_g \longrightarrow X\,^3\Sigma_g^-$. The nightglow Atmospheric ($b \longrightarrow X$) 0–1 band has an emission rate of $4\pi\mathcal{I} \approx 1{\cdot}5$ kR (Dufay[24]; Barbier[25]). The photon emission in the 0–0 band of the same system (which is absorbed by the lower atmosphere) is 21 times that in the 0–1 band, according to the Franck–Condon factors computed by Fraser, *et al.*[14]. The total zenith emission is then $4\pi\mathcal{I}_{b\Sigma} \approx 30$ kR in the Atmospheric system. In the Infrared Atmospheric bands the nightglow emission should be produced by direct excitation as well as by cascading $b \longrightarrow a$.

Hence in the absence of electronic deactivation, we should have $4\pi \mathcal{J}_{aX} \geq (A_{ba}/A_{bX})$ 30 kR. From the twilight decay of the $a \rightarrow X$ (0–1) band at 1·58μ, measured by Vallance Jones and Harrison[26], it appears that its nightglow emission is no more than 5 kR. The unobservable 0–0 band would emit 10 times as strongly, so we conclude that $4\pi \mathcal{J}_{aX} \leq 50$ kR. Hence the observations indicate that $A_{ba}/A_{bX} \leq 5/3$ or that $A_{ba} \leq 0{\cdot}20$ sec^{-1}. This value is too low to increase materially the efficiency of reaction (1*a*) by removing O_2 ($b\,^1\Sigma_g^+$) molecules.*

Seaton's[5] computation of the deactivation probability of O (^1D) at various heights used the following general approach. Observations of the intensity ratio of the red to green line versus height were interpreted in terms of the deactivation required at each height. Then adopting $A_{ba} + A_{bX} = 1$ sec^{-1}, he adjusted the rate coefficients for reaction (1*a*) and for an unspecified deactivation process proportional to the gas density to fit a theoretical curve to the empirical points. There is considerable uncertainty, however, in estimating an empirical deactivation rate versus altitude because of the lack of information on the *excitation* rates of the red and green lines.

With the additional but crude information on deactivation of O (^1D) by O_2 collisions obtained in Section 3, we shall here consider the deactivation that would be expected from reactions (1*a*), (1*b*) and (2) at various heights.

To write the equation of statistical equilibrium for O (^1D), we assume a rate of excitation per cm^3 per sec, from all processes *except* energy transfer from O_2 by reaction (1*b*), to be q_D. Then, with additional excitation by (1*b*) and deactivation by (1*a*) we have

$$q_D + [\mathrm{O}]\,[\mathrm{O_2}^*(2)]\,\alpha_{1b} = [\mathrm{O}^*]\left\{[\mathrm{O_2}]\sum_{i=0}^{2}\alpha_{1a}(i) + A_{D'}\right\}. \quad (11)$$

In order to apply the results to the airglow, where O_2* is formed by processes other than (1*a*), let us assume that three-body association populates $b\,^1\Sigma_g^+$, $v' = 2$ with a rate $q_2(2) = [\mathrm{O}]^2\,[X]\,\beta(2)/\mathrm{cm}^3/\mathrm{sec}$. Then equation (4) is modified to read

$$[\mathrm{O_2}^*(2)] = \frac{q_2(2) + [\mathrm{O_2}]\,[\mathrm{O}^*]\,\alpha_{1a}(2)}{[\mathrm{O}]\,\alpha_{1b} + [\mathrm{O_2}]\,\alpha_2 + A_{bX}}. \quad (12)$$

Substituting equation (12) into (11) we may write the result as

$$q_D + \frac{q_2(2)\,[\mathrm{O}]\,\alpha_{1b}}{[\mathrm{O}]\,\alpha_{1b} + [\mathrm{O_2}]\,\alpha_2 + A_{bX}} = [\mathrm{O}^*]\,(d_D + A_{DP}), \quad (13)$$

where the deactivation probability of O (^1D) is

$$d_D = [\mathrm{O_2}]\sum_{i=0}^{2}\alpha_{1a}(i) - \frac{[\mathrm{O_2}]\,\alpha_{1a}(2)\,[\mathrm{O}]\,\alpha_{1b}}{[\mathrm{O}]\,\alpha_{1b} + [\mathrm{O_2}]\,\alpha_2 + A_{bX}}. \quad (14)$$

In equation (13) the second term on the left is the fraction of the primary excitation of O_2* (2) that eventually is transferred to O (^1D). Equation (14) gives the net deactivation probability to be the probability that an O (^1D) atom will transfer its excitation by reaction (1*a*) *less* the fraction of those entering reaction (1*a*) that are eventually returned to O (^1D) by (1*b*). But with the low value of the rate coefficients $\alpha_{1a}(2)$ and α_{1b} indicated in Table 2, the second term in equation (14) is negligible at all heights in question.

Thus it appears that deactivation of the red lines is nearly proportional to $[O_2]$. The effective rate coefficient for deactivation, $\alpha_{DP} = \Sigma\alpha_{1a}(v)$, is greater than* 4×10^{-12} cm^3/sec, according to the estimates of Table 2. A coefficient $\alpha_{DP} \sim 10^{-11}$ cm^3/sec would, according to Seaton, satisfy the observations at altitudes of 130 km and higher. At lower

* If the fluorescence mechanism (discussed in Section 3) for converting 0–0 emission into the 0–1 band, were not inhibited from proceeding by deactivating collisions, the total emission would be as low as 8 kR. But if fluorescence did affect the total derived intensity of $b \rightarrow X$ it would also affect the $a \rightarrow X$ system and diminish our estimate of the total emission in those bands. Hence the above argument against a large value for A_{ba} would not be significantly altered.

* If the rate coefficients $\alpha_{1a}(v)$ should decrease rather rapidly with increasing v, as do the Franck–Condon factors in the $v'' = 0$ progression of the Atmospheric system, α_{DP} may be considerably greater.

altitudes the coefficient might be considerably less. As we mentioned previously, however, the excitation rate and its variation with height is so uncertain that an empirical deactivation rate is rather questionable. For the moment a coefficient for deactivation by O_2 of $\alpha_{pD} \sim 10^{-11}$ cm^3/sec seems acceptable and reasonably consistent with all the data.

We are indebted to Prof. D. R. Bates for commenting on an early draft and for several illuminating discussions.

REFERENCES

1. A. B. Meinel, *Astrophys. J.* **112**, 464 (1950).
2. A. B. Meinel, *Astrophys. J.* **113**, 583 (1951).
3. G. Kvifte, *Nature, Lond.* **168**, 741 (1951).
4. J. W. Chamberlain, C. Y. Fan and A. B. Meinel, *Astrophys. J.* **120**, 560 (1954).
5. M. J. Seaton, *Astrophys. J.* **127**, 67 (1958).
6. J. W. Chamberlain, *Astrophys. J.* **121**, 277 (1955).
7. H. D. Babcock and L. Herzberg, *Astrophys. J.* **108**, 167 (1948).
8. G. Herzberg, *Spectra of Diatomic Molecules*, p. 127. Van Nostrand, Princeton, N.J. (1957).
9. R. J. Schlapp, *Phys. Rev.* **51**, 343 (1937).
10. J. Dufay and M. Dufay, *C. R. Acad. Sci., Paris* **232**, 426 (1951).
11. C. E. Moore, *A Multiple Table of Astrophysical Interest. Contr. Princeton Univ. Obs.* No. 20 (1945).
12. J. W. Chamberlain and A. B. Meinel, *The Earth as a Planet* (Edited by G. P. Kuiper), Chap. 11. University of Chicago Press (1954).
13. R. C. Herman and G. A. Hornbeck, *Astrophys. J.* **118**, 214 (1953).
14. P. A. Fraser, W. R. Jarmain and R. W. Nicholls, *Astrophys. J.* **119**, 286 (1954).
15. R. W. Nicholls, *Ann. Géophys.* **14**, 208 (1958).
16. D. R. Bates and A. Dalgarno, *J. Atmos. Terr. Phys.* **4**, 112 (1953).
17. M. J. Seaton, *J. Atmos. Terr. Phys.* **4**, 295 (1954).
18. D. R. Bates, *The Earth as a Planet* (Edited by G. P. Kuiper), Chap. 12. University of Chicago Press (1954).
19. D. R. Bates, *Proc. Phys. Soc. London* **A68**, 344 (1955).
20. H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena*. Clarendon Press, Oxford (1952).
21. J. W. Chamberlain, *Astrophys. J.* **119**, 328 (1954).
22. M. Nicolet, *Ann. Géophys.* **15**, 1 (1959).
23. A. Omholt, *J. Atmos. Terr. Phys.* **10**, 320 (1957).
24. M. Dufay, *C. R. Acad. Sci., Paris* **246**, 2281 (1958).
25. D. Barbier, *Vistas in Astronomy* (Edited by A. Beer), Vol. II. Pergamon Press, London (1956).
26. A. Vallance Jones and A. W. Harrison, *J. Atmos. Terr. Phys.* **13**, 45 (1958).

Planet. Space Sci. Pergamon Press 1959. Vol. 2, pp. 71–77. Printed in Great Britain.

RESEARCH NOTES

Relative intensities of Na lines in the emission of sodium ejected from rockets

(*Received* 7 *July* 1959)

In 1950, D. R. Bates[1] detailed the possibilities offered by an artificial sodium cloud to get information concerning the high atmosphere. Since that time, experiments have been performed in the United States by Manring, Bedinger and others[2,3,4]. Recently we have undertaken a series of experiments with French missiles with a view to study the emission of the ultraviolet doublet 3303 ($4P$–$3S$) and the intensity ratio of the lines 5893 and 3303.

The first experiments were made during last March in French Sahara, where two "Véronique" rockets were fired, the first at 1840 hr GMT on 10 March and the second at 0544 hr GMT on 12 March; the first reached an altitude of 126 km and the second 180 km.

The experimental details have been given in a previous publication[5]. Since our problem was the comparison of the intensities of the two lines, yellow and ultraviolet, it was of the utmost importance to get the two lines on the same spectrum. Hence the spectrograph was adjusted for focus in the two spectral regions and we used Kodak 0aF plates, sensitive from 6800 Å to the ultraviolet. One spectrum was obtained for each firing; on 10 March the exposure time was 30 min, and on 12 March 15 min only, owing to the increasing light of the dawn.

The plates and spectrograph were then calibrated at the laboratory with a Philips sodium lamp which emits both lines with an intensity ratio which has been carefully measured. We found $I_{5893}/I_{3303}=1277$.

The first spectrum, obtained in the evening twilight, gives no evidence of the 3303 line; the second, at the morning twilight, gives a weak line at 3303; we must note that the exposure time was shorter and that the yellow line was less than half the intensity obtained in the first case. Hence we may immediately deduce that the intensity ratio of these two lines is variable, but so far we are not in a position to say whether the difference comes from the different altitudes, or from the time of firing or from the atmospheric conditions.

In spite of the difficulty of the comparison, due to the weakness of the ultraviolet line, we could determine the ratio of the energy intensities received at ground level: the intensity of the yellow line was 1–5 times the intensity of the ultraviolet line. We are conscious of the unavoidable sources of error of this measurement, and the figures given above are only a rough estimate of the intensity ratio; but it is of interest to compare this value to the similar measurement made by Cooper *et al.*[3], who have found a ratio of more than 1:100. We find that, in our experiment, the ultraviolet line was far more intense than 1/100th of the yellow line, while the ratio 1:100 is in agreement with the theory of resonance scattering. A ratio of, say, 1:2 or 1:3 would indicate a quite different excitation. This means that different mechanisms are simultanously or successively present in the mesosphere.

Further, if we contemplate the fact that in the first experiment 3303 line was not visible, we are brought to the conclusion that further investigations are necessary. We must remember that Bedinger *et al.*[4] have not observed the ultraviolet line at night nor in the evening, and also that it is in the evening that the 3303 line is missing in our own experiments.

In further experiments we intend to replace the ultraviolet spectrograph by a quartz spectrograph, to increase the transmission for 3303 line and to vary the conditions of the firing in order to examine the possibility of different kinds of excitation of the sodium ejected from the rockets. We intend also to find the spatial region where 3303 line is excited.

Such information may be of importance for a better knowledge of the state of the upper atmosphere, which is substantially variable with altitude as was already noted in the U.S. experiments.

A. AND E. VASSY

Physique de l'Atmosphère
Faculté des Sciences de Paris

References

1. D. R. BATES, *J. Geophys. Res.* **55**, 347 (1950).
2. J. F. BEDINGER and E. MANRING, *J. Geophsy. Res.* **62**, 170 (1957).
3. C. D. COOPER, E. MANRING and J. F. BEDINGER, *J. Geophys. Res.* **63**, 369 (1958).
4. J. F. BEDINGER, E. R. MANRING and S. N. GHOSH, *J. Geophys. Res.* **63**, 19 (1958).
5. A. and E. VASSY, *C.R. Acad. Sci., Paris* **248**, 2235 (1959).

The current-jet hypothesis of whistler generation

(*Received* 10 *August* 1959)

The paper of H. Norinder and E. Knudsen, on "The relations between lightning discharges and whistlers"[1], is of considerable interest from the standpoint of generation of whistlers. This part of the phenomenon has received little attention up to now, apart from the fact that the whistler is associated in time with a spheric, though it is known that not every spheric gives rise to a whistler. Only the high-energy sferics appear to be capable of generating whistlers, and of these only those with "slow tails" (which, as Norinder points out in his paper, apparently correspond to multiple discharges). The time lags involved in the "slow tails" are of the order of tens or hundreds of milliseconds, and the spheric energy peaks at around 5 kc/s, rather than near 10 kc/s as in the ordinary strong spheric.

The insertion and passage of the extraordinary component of the spheric signal through the lower ionosphere, with the high values of collision frequency encountered there, appears to be the least understood part of whistler propagation. It is difficult to understand how upon occasions, long echo trains of twenty to thirty round trips can be achieved under such circumstances, nor how occasional amplification of whistler echoes can occur. (This phenomenon has been attributed to "polar creep", but either polar or equatorial creeps would be inconsistent with the widely accepted current belief in ducts of field-aligned ionization.) The diurnal variation in frequency of occurrence of whistlers is also consistent with strong absorption in the lower ionosphere.

It is the purpose of this letter to outline what might be termed the current-jet theory of whistler generation and to review the evidence for such a hypothesis. In this theory, the penetration of the lower ionosphere is accomplished by a high-energy current jet which is emitted from the top of the thundercloud shortly after the spheric discharge occurs. Until it enters the ionosphere, the current jet is accelerated by the electrostatic field of the thunderstorm, and also interacts with one or more of the slow-tail portions of the spheric in such a way as to either amplify the spheric signal or be accelerated by it. During the passage through the lower part of the ionosphere the current jet becomes field-aligned and undergoes synchronous acceleration by a later spheric. During this phase, additional acceleration along the field line takes place due to the acceleration by collisions at the "magnetic mirrors" at each end of the path. Meanwhile a travelling-wave-type interaction is taking place between the high-energy electrons launched in the current jet and the electromagnetic wave of the spheric signal. These high-energy electrons are trapped in the geomagnetic field and are reflected back and forth by the "magnetic mirror" in the lower ionosphere. The result is that the thunderstorm creates its own duct of field-aligned ionization, which guides and interacts with the spheric signal either to amplify or damp it out depending upon the relative velocities of electrons and electromagnetic waves. Such accelerated current-jet electrons could also contribute in thunderstorm latitudes to the inner radiation belt, a slow tail of 1 sec duration corresponding to energies (for electrons) of 20 keV or more.

Let us now review the evidence for such a hypothesis. Although the existence of upward

discharges from thunderclouds fell into disrepute after some negative results were obtained in the late 1930's, the phenomenon seems to be well substantiated[2] and has been personally witnessed by the writer. Perhaps the most significant evidence on the matter is the radar observations of G. C. Rumi[3] at 27·85 Mc/s. Rumi observed a number of discrete columnar echoes, which travelled upward in the sub-ionospheric region at speeds up to 200 km/sec. He attributed these to upward discharges. His paper, incidentally, contains an excellent summary of the previous work on this subject, including C. T. R. Wilson's[4] theory of the vertical acceleration of the upgoing electrons by the electric moment of the thundercloud and the work of Watson Watt, Wilkins and Bowen[5] on the curvature of the upward discharges in the geomagnetic field. One notes that in the case considered by the latter authors, only the high-energy electrons (applied voltage$>$1·5 x 10^{8}V) had sufficient energy to reach the ionosphere, the others either returning to earth or being collected along the conducting strata of the stratosphere in the neighbourhood of the ozonosphere or the conducting layer at about 60-km altitude observed by Seddon in rocket flights. We note that the latter altitude (60 km), regarded as a wavelength, corresponds to a frequency of 5 kc/s, which also happens to be the frequency at which the energy spectrum of whistler-producing spheric peaks up, while the ozonosphere, centered at about 30-km altitude, corresponds to a frequency of 10 kc/s. As it happens, this is the frequency at which the energy spectrum of non-whistler producing spheric peaks up.

Wilson estimated that electrons would be accelerated to energies of several tens of kilovolts by the electrostatic field of the cloud. He further noted that a conducting atmospheric layer at an altitude of 60 or 80 km increases the vertical electrical force below it. If the field exceeds a certain critical value, the ions of the conducting layer acquire sufficient energy to cause ionization by collision. If the field is of sufficient strength to bring about collisional detachment for a considerable distance below the conducting layer, discharge of the upper pole of the thundercloud to the upper atmosphere can occur. The electron velocities are thus often of the right order (0·1 c) for synchronous interaction with the *r-f* wave. In the ionosphere and exosphere the particles are further accelerated by the "magnetic mirror" effect. The full theory of the accelerative process in detail remains to be developed, but portions already exist in magnetohydrodynamics and the theory of gaseous discharges, and the rest is discernible in its essential features.

The validity of the proposed hypothesis of whistler generation can be tested quite simply by making simultaneous observations with a whistler receiver and an *Hf* radar equipment similar to that described by Rumi. Time coincidence (with appropriate lags) of whistler signals and echoes from upward discharges would provide a powerful argument for the current-jet hypothesis.

W. C. HOFFMAN

Hughes Research Laboratories
Culver City, California

References

1. H. NORINDER and E. KNUDSEN, *Planet. Space Sci.* **1**, 173 (1959).
2. J. A. CHALMERS, *Atmospheric Electricity*, p. 322. Pergamon Press, London (1957).
3. G. C. RUMI, *J. Geophys. Res.* **62**, 547 (1957).
4. C. T. R. WILSON, *Proc. Phys. Soc.* **37**, 32D (1925).
5. R. A. WATSON WATT, A. F. WILKINS and E. G. BOWEN, *Proc. Roy. Soc.* **A161**, 181 (1937).

On the polarization of the oxygen red line in aurorae*

(*Received* 16 *September* 1959)

In a recent paper Duncan[1] reported a 30 per cent polarization of λ 6300 [O I] in a low-latitude, red aurora. The interpretation of this

* The research reported in this paper was supported in part by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19(604)-3044 with the University of Chicago.

result was that the exciting electrons possess a highly preferential motion. Duncan gave plausibility arguments, based on semi-classical theory, that this motion would be perpendicular to the magnetic field (i.e. a gyrational motion), rather than along it. A correct explanation of the measurement is not, however, immediately apparent, and, in fact, it is not at all clear how such a large polarization could arise.

Polarization is a direct consequence of the Zeeman effect and a precise calculation of the polarization requires a knowledge of how the magnetic quantum number M changes upon excitation—i.e. how the angular momentum of the atom changes direction. It is difficult to make any qualitative statements from the semi-classical theory, which is notorious for predicting incorrect polarizations, especially when there is electron exchange, as must be the case in excitation of the ^3P–^1D multiplet.

The magnetic field not only might be responsible for preferred directions of motion of exciting particles, but for a forbidden line, such as λ 6300, it also produces magnetic depolarization, because degeneracy is fully removed. The separation of the Zeeman states ($\Delta\nu \sim eH/4\pi mc \sim 10^6$ sec^{-1}) is large compared with the natural width ($A/2\pi \sim 10^{-3}$ sec^{-1}); or the atom will gyrate many times during its lifetime, insuring axial symmetry about the field. For exciting electrons whose motions are symmetric about the field, magnetic depolarization does not introduce any new effect on the angular dependence of the polarization. It is mentioned here to emphasize that there is symmetry about the field, and depolarization for observations along the field, regardless of the excitation mechanism —e.g., by resonance scattering in twilight.

The red lines arise predominantly from magnetic-dipole transitions, which are entirely analogous to electric-dipole transitions if the electric and magnetic vectors, as well as the dipoles, are interchanged throughout (see Condon and Shortley[2] p. 93). Let A_π^M be the probability for a transition in the line λ 6300 ($^3P_2 - {}^1D_2$) from Zeeman state M in the upper ($^1D_2$) level to the lower level, with emission of a component polarized with the magnetic vector parallel to the external field. Similarly, A_σ^M is the total probability for emission of components polarized perpendicular to the field. These partial transition probabilities are readily obtained from known intensity rules (see Condon and Shortley[2], p. 387). The polarization may then be expressed as

$$P(\Theta) = \frac{\mathscr{I}_\perp - \mathscr{I}_{||}}{\mathscr{I}_\perp + \mathscr{I}_{||}} = \frac{\beta \sin^2 \Theta}{\alpha - \beta \sin^2 \Theta}. \qquad (1)$$

Here Θ is the angle between the magnetic field and the line of sight, and $\mathscr{I}_\perp$ and $\mathscr{I}_{||}$ are the intensities with the magnetic vector perpendicular and parallel, respectively, to the plane formed by the line of sight and the magnetic line of force through the emitting atom. The quantities α and β are

$$\alpha = \Sigma_M N_M A_\sigma^M = A\left(\frac{2}{3}N_2 + \frac{5}{3}N_1 + N_0\right) \qquad (2)$$

and

$$\beta = -\Sigma_M N_M \left(A_\sigma^M - \frac{1}{2}A_\sigma^M\right) = A\left(-N_2 + \frac{1}{2}N_1 + \frac{1}{2}N_0\right) \qquad (3)$$

where A is the total transition probability for λ 6300, N_M is the population of state M, and the numerical values assume $N_{-M} = N_M$.

The largest possible *negative* polarization (magnetic vector north-south for zenith observations) is obtained by setting $N_1 = N_0 = 0$; then $\beta/\alpha = -3/2$. Equation (1) then gives $P_{\max}(90°) = -3/5$. However, Duncan's observations were made at geomagnetic latitude 42°S, where the field lines are inclined to the vertical by approximately 30°. Thus the maximum polarization for zenith observations would be $P(30°) = -3/11$, which is about the polarization reported.

There are a number of effects, however, that will lower the actual polarization from this theoretical maximum: (1) A preferential direction in the motion of the exciting electrons (for excitation by inelastic collisions) would be expected to produce some polarization, but the relatively slow electrons important for exchange collisions should have a nearly isotropic

distribution. Thus only a small polarization would have been anticipated, whether the electrons are primaries that have suffered many decelerating collisions, ejected secondaries, or electrons locally accelerated by the action of hydromagnetic waves* or even electric fields. (2) Even with a highly directed beam, any inelastic collisions that excite $M=0$ or ± 1 will decrease the polarization. (3) Depolarizing collisions, which change the Zeeman state of the upper level (i.e. change the orientation of the excited atom), may be important for a level with such a long lifetime.

The need for polarization measurements on the aurora was emphasized earlier by Ginsburg[3], who showed, incidentally, that the green line would not be polarised.† He suggested that molecular nitrogen bands, which may be excited in part directly by incident primaries, would be the most likely emissions to be polarized. A high polarization for λ 6300 is quite startling and would scarcely have been predicted.

J. W. CHAMBERLAIN

Yerkes Observatory
University of Chicago
Williams Bay, Wisconsin

References

1. R. A. DUNCAN, *Planet. Space Sci.* **1**, 112 (1959).
2. E. U. CONDON and G. H. SHORTLEY, *The Theory of Atomic Spectra*. Cambridge University Press (1951).
3. V. L. GINSBURG, *Doklady Akad. Nauk SSSR* **38**, 237 (1943).

* This process, which may be regarded essentially as a means of raising the electron temperature, is attractive insofar as the low-latitude aurorae, with their extremely high red/green intensity ratios, are concerned. As in the electric-discharge mechanism, the excitation of the red lines could be enormously greater than for λ5577, whereas impact by fast, extraterrestrial primaries and their secondaries presents some difficulties in regard to this ratio.

† Ginsburg's calculation of the polarization of λ6300 from resonance scattering (2/79) used quadrupole transition probabilities for the Zeeman components. The transition is, however, principally magnetic dipole, and dipole components give $\beta/\alpha = 7/85$ or $P(90°)=7/51$ for no magnetic field or a magnetic field parallel to the incident beam. For other orientations there is no difficulty in principle in computing the polarization.

Micrometeorite impacts to an altitude of 135 km

(*Received 5 October* 1959)

Long term measurements of the average influx rate of micrometeorite particles (diameters about 10μ) or larger meteoritic particles have been made in artificial earth satellite vehicles using ballistic piezoelectric pickups[1] and grid abrasion detectors[2]. For both methods, average counting rates are probably about $1{\cdot}7 \times 10^{-3}$ impacts/m²/sec. Dubin (reported by Manring[2]) used microphone detectors and found that the average influx rate for particles with diameters greater than about 3μ was around 3×10^{-2} impacts/m²/sec. Superimposed on this low background influx Komissarov and others[1] found short term influx rates as high as 22 impacts/m²/sec between 1700 and 1880 km, with others of 10 impacts/m²/sec between 1300 and 1500 km and 9 impacts/m²/sec between 500 and 600 km suggesting the existence of micrometeorite streams which reach the earth from different directions. Similarly Richter (reported by Kellogg[3]) has suggested that the U.S. satellite Explorer III may have been hit by the eta-Aquarid meteor shower when two of the wire detectors were severed almost simultaneously, followed by erratic behaviour in the high-power transmitter.

In direct contrast to the low average infall rates measured in satellites, Berg and Meredith[4] reported a very much higher impact rate on a detector in an Aerobee rocket fired to 103 km. Their detector was a polished lucite cone set in a plastic housing and mounted with its flat base recessed into the outer skin of the rocket and its apex directed inside the rocket. The lucite cone was completely covered with a thin (8×10^{-6} cm) opaque layer of aluminium except for a clear exit window at its apex. A photomultiplier was placed adjacent to this window and the small light flashes emitted by the impacting micrometeorites which penetrated the opaque aluminium coating of the detecting surface was captured within the lucite cone and reflected into the photomultiplier. The output pulses from the photomultiplier were suitably

amplified, stretched and then telemetered back to a ground recording station. Berg and Meredith estimated that their detector would detect impacts of micrometeorites having energies as low as 0·005 erg, corresponding to an iron micrometeorite with a diameter of 1μ and a velocity of only 0·5 km/sec. The result of their experiment indicated very few impacts below 85 km and an approximately constant, average rate of about $1{\cdot}6 \times 10^2$ impacts/m^2/sec above this altitude. This rate is about five orders of magnitude greater than the satellite rates.

In view of this discrepancy, it was decided to repeat the Berg and Meredith experiment using a Long Tom two-stage rocket fired from the Woomera Rocket Range, Australia, as part of the IGY Upper Atmosphere Research Program. Dr. Berg very kindly supplied constructional details of the detector and amplifier so that the present experiment was essentially identical to that performed by Berg and Meredith except in the following minor respects. In view of the fact that most of the aluminized surface of one of the detectors used by Berg and Meredith had been burned off in the denser parts of the atmosphere, it was decided to cover the exposed detector surface in the present experiment with a metal band which was automatically removed at an altitude of 53 km. Since the Long Tom rocket reaches accelerations of up to 70g, special precautions had to be taken to ensure that the photomultiplier and the vacuum tubes in the amplifier would not be damaged during the firing. The photomultiplier tube was accelerated to 70g for a few seconds on a centrifuge and X-ray examination before and after acceleration showed no change in internal structure and the tube operated perfectly. The performance of the amplifier over the duration of flight of the round was monitored by feeding a test pulse into the amplifier. This test pulse was recorded at the ground station for the whole of the time of flight of the round and indicated that the amplifier was operating satisfactorily throughout.

The round (*LT* 10) was fired at night (1022 hours GMT) on May 27th, 1959 and reached an altitude of 135 km. The duration of flight was 495 sec, the cover was removed from the detector at 53 km after 50 sec of flight and 210 sec were spent at altitudes above 85 km. If the influx of micrometeorites is as high as Berg and Meredith reported, then about 276 impacts should have been recorded in this time (detector area 82 cm^2); however no impacts were recorded during the entire duration of the experiment. There is no reason to assume that the photomultiplier failed during the experiment and the amplifier was shown to be operating throughout, so that the absence of recorded impacts does not seem to be attributable to equipment failure.

Since no impacts were recorded, then no statement can be made about average influx rates, however it is possible to give an upper limit to the influx rate.

If the influx rate (λ) is assumed to be constant, then the probability that an impact will be recorded in any short time interval (dt) on a given area (A) is given by $A\lambda dt$. Furthermore, if the probability of two impacts occurring within time dt is an infinitesimal of a higher order, the probability (P) of recording X hits in time t on an area A is given by:

$$P = \frac{(A\lambda t)^X}{X!} \cdot \exp^{(-A\lambda t)}$$

Since $A = 82 \times 10^{-4}$ m^2 and $t = 210$ sec, it is then possible to allocate the following upper bounds to the influx rate corresponding to various confidence coefficients:

Confidence coefficient	Upper bound for λ (impacts/m^2/sec)
0·95	2·1
0·99	3·1
0·999	4·2

Thus the probability is $\geqslant 0{\cdot}95$ that the upper bound of 2·1 impacts/m^2/sec is greater than the true impact rate of micrometeorites.

This upper limit includes the satellite rates and so is consistent with them, however it is about two orders of magnitude less than the average rate calculated by Berg and Meredith.

The reason for this discrepancy is not certain but it may be that their high influx rates may be explained by the fact that their count was made during the Leonid meteor shower whereas the time at which the present experiment was made did not coincide with any known meteor shower. If this explanation is correct, then the influx rate of meteor particles during the occurrence of meteor streams may be as high as five orders of magnitude greater than the background influx rate of micrometeorites given by the satellite experiments.

Acknowledgements—The success of this experiment is due to the co-operation of many individuals—in particular Professor L. G. Huxley and E. L. Murray of Physics Department, Adelaide University, Dr. A. R. Hogg, G. R. Graf and W. van Barneveld, Australian National University, and the "Long Tom" development group, Weapons Research Establishment, Salisbury, South Australia. Statistical problems were discussed with Professor E. J. Hannan, Canberra University College.

J. F. LOVERING

Department of Geophysics
Australian National University
Canberra

References

1. O. D. KOMISSAROV, T. M. NAZAROVA, L. N. NEUGODOV, S. M. POLOSKOV and L. Z. RUSAKOV, *Artificial Earth Satellites, Acad. Sci. SSSR*, No. 2, 54 (1958).
2. E. R. MANRING, *Planet. Space Sci.* **1**, 27 (1959).
3. W. W. KELLOG, *Planet. Space Sci.* **1**, 71 (1959).
4. O. E. BERG and L. H. MEREDITH, *J. Geophys. Res.* **61**, 751 (1956).

VOL
2
1959

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 79–85. Printed in Great Britain.

PHOTOCHEMICAL FORMATION OF ORGANIC COMPOUNDS FROM MIXTURES OF SIMPLE GASES

W. E. GROTH and H. v. WEYSSENHOFF
Institut für Physikalische Chemie, Universität Bonn

(*Received* 23 *June* 1959)

Abstract—Experiments are described on photochemical reactions in mixtures of methane or ethane with ammonia and water vapour which were irradiated with light of extremely short wavelengths. Both mercury-sensitized and direct photolysis experiments were carried out. Sensitized experiments with a mixture of methane, ammonia and water vapour did not yield any reaction products. Amino acids, low fatty acids and liquid hydrocarbons were obtained when methane was replaced by ethane. Amino acids and fatty acids were quantitatively determined. In direct photolysis experiments with mixtures of methane, ammonia and water vapour, using the resonance radiation of xenon, glycine was formed which was detected by paper chromatography. The relation of these experiments to the photochemical processes in the atmosphere of the earth in its early stages is briefly discussed.

1. INTRODUCTION

As hypothesized by A. I. Oparin[1], H. C. Urey[2], W. Kuhn[3] and J. D. Bernal[4], the earth's atmosphere was in a reduced state in the early stages of its evolution. This atmosphere contained hydrogen, methane, ammonia and water vapour but no free oxygen. Under the influence of ultraviolet sunlight as well as electric spark discharges and lightning these gases were dissociated. While hydrogen escaped into space[3] the remaining gases were slowly oxidized. In the course of this oxidation process many simple organic compounds were possibly formed in the atmosphere as intermediate products. These compounds were washed down with moisture and accumulated in the oceans, thus providing favourable conditions for the evolution of primitive living organisms[1, 4, 5].

The formation of organic compounds under conditions which are present in the reducing atmosphere was demonstrated by various laboratory experiments. S. L. Miller[6] as well as K. Heyns *et al.*[7] and H. A. Rathsack[8], exposed a mixture of methane, ammonia, water vapour and hydrogen to an electric spark discharge. From these experiments they obtained several amino acids, carboxylic acids, amines and other organic compounds.

Electric discharges were probably not the strongest source of energy although they would form ions and radicals of high concentration near the earth's surface. As emphasized by Urey[2] and Miller[6], the greater part of the energy would be contributed by the ultraviolet radiation of the sun.

For this reason, gas mixtures consisting of methane or ethane in combination with ammonia and water vapour were irradiated with ultraviolet light in mercury sensitized as well as direct photolysis experiments. For irradiation the resonance wavelengths at 2537 Å and 1849 Å of mercury, and at 1470 Å and 1296 Å of xenon were used. Some results of these experiments were reported previously[9]. A detailed description of the experimental technique and a discussion of the results are given in the present paper.

2. APPARATUS AND ANALYTICAL METHODS

For the mercury-sensitized experiments, a circulating system was used consisting of a mercury low-pressure discharge lamp, a

mechanical circulating pump[10], a 5 liter flask, a small boiling flask and a condenser, as shown in Fig. 1. The mercury quartz lamp has the shape of a Liebig condenser with a length of 1 m. Its quantum yield was determined by measuring the photodissociation of monochloroacetic acid[11] to be 3×10^{18} quanta/sec.

The gas mixture circulates in the direction marked by the arrows in Fig. 1. At *D*, water vapour enters the gas circulation path and is condensed, after leaving the lamp, in the condenser together with the reaction products. The condensate is led back through the *U*-tube into the boiling flask. The flask is filled with 100 ml of bidistilled water through the side arm *A* which also serves for the removal of the product solution for analysis. The tube connecting the flask with the lamp is heated by an asbestos heating coil in order to prevent condensation of the water.

Before starting each run the whole apparatus was carefully washed by means of alkaline permanganate solution. For all runs with sensitizer, a droplet of mercury was heated in the evacuated jacket of the lamp.

The sensitized experiments using mixtures of methane, ammonia and water vapour did not yield any reaction products, since methane does not absorb appreciably above 1450 Å[20]. For this reason, a radiation of wavelengths shorter than 1450 Å had to be applied. As a light source, a quartz lamp was used of the type developed by P. Harteck and R. Oppenheimer[12], and W. Groth[13]. It was operated with a discharge in a neon-xenon mixture at ca. 50 A which emitted the xenon resonance wavelengths at 1470 and 1296 Å through a calcium fluoride window. The neon in the discharge had to be purified by extensive circulation over cooled charcoal traps before xenon was added. The top of the lamp was located within a 3000 G magnetic field in order to concentrate the discharge directly in front of the window. The lamp emitted a maximum quantum yield of 2×10^{16} quanta/sec which was determined by photochemical formation of ozone.

It has been attempted to operate the lamp with a neon-krypton discharge using lithium fluoride windows. In this case the resonance

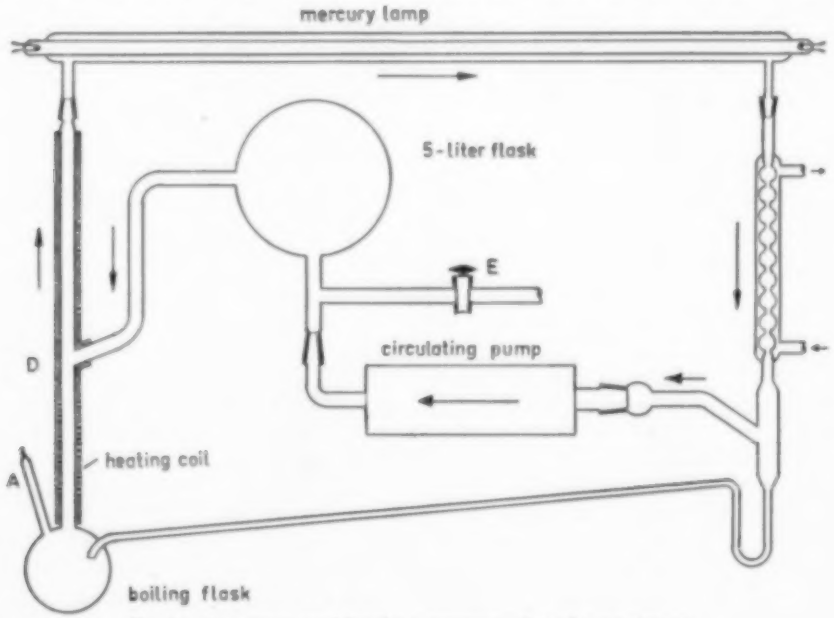


Fig. 1. Apparatus used for the mercury-sensitized experiments.

wavelengths at 1236 and 1165 Å would be available which had the advantage to be more efficient for the photodissociation of methane. The transparency of the lithium fluoride windows was, however, reduced by a factor of 100 during each run whereas that of the calcium fluoride windows remained nearly constant.

A thermal convection circulating system was used as reaction vessel as is shown in Fig. 2.

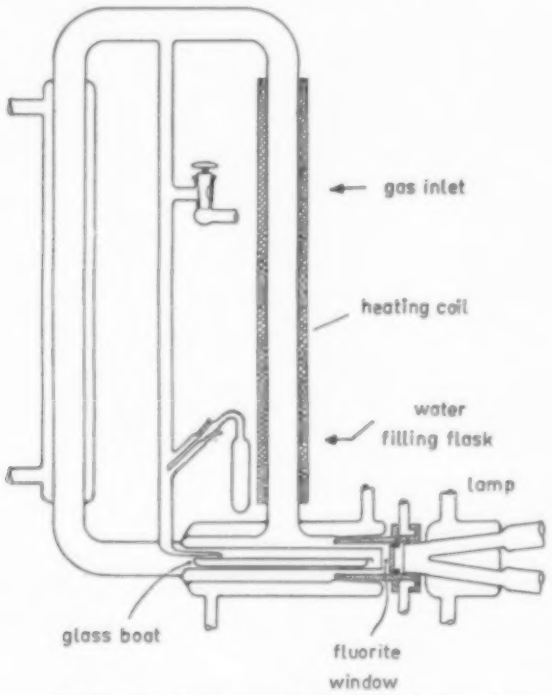


Fig. 2. Circulating system operated with thermal convection.

The right arm of the system is heated to 200°C. The left arm and the reaction chamber are cooled with water which is kept at constant temperature by a thermostat. A little glass boat which contains water is situated directly behind the reaction zone. Thus the gas mixture and any reaction products come into contact with the water immediately after irradiation. The water is supplied to the boat by rotating the small flask into which water had been distilled before admitting the gases. The apparatus is connected to the lamp by a metal mounting to which the window is cemented as well. The mounting is cooled by the same thermostated water as the other parts of the apparatus.

For reasons discussed below, the temperature of the cooling water had to be kept at 55°C. By means of heating coils all free parts of the circulating system were kept at this temperature in order to prevent the water inside the glass boat from condensing elsewhere. The temperature was limited by the properties of the cements used for mounting the window and connecting the apparatus with the lamp.

At elevated temperatures the most reliable material for cementing the window was a polymerized resin* which, however, contains nitrogen bonds. It was confirmed by control

* *Araldit*, produced by Ciba, Basel.

tests that traces of this resin found their way into the gas mixture and thus obscured the detection of photochemical reaction products. This source of error could be removed by a teflon seal which consists of a ring-shape cut from a thin sheet. This ring was pressed tightly against the edge of the window by means of a threaded cylinder. The arrangement is shown in Fig. 3.

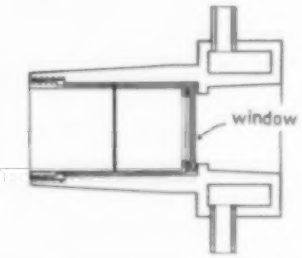


Fig. 3. Arrangement of the Teflon seal

Preparation of the gases

Methane, ethane and ammonia were available with a purity of at least 99 per cent. The gases were distilled several times, each middle fraction being retained. During this operation it was taken care that the gases did not come into contact with mercury.

Analytical methods

The solutions resulting from the experiments with the mercury lamp were analysed by a procedure similar to that used by Miller[6].

Amino acids were detected by a paper chromatographic test. The chromatograms were developed by a descending technique using a mixture of 50 per cent water, 40 per cent *n*-butanol and 10 per cent acetic acid, or water saturated phenol.

The original solution containing the reaction products and ammonia absorbed from the gas phase was separated into two fractions by passing through a cation exchange column of 30 cm length and 12 mm in diameter, filled with Dowex 50 in H^+-form. Acids and non-ionic compounds were obtained by elution with water. Amino acids were eluted by a 2 N solution of ammonia. Basic compounds stronger than ammonia, e.g. simple amines, were retained in the column. These compounds were eluted with hydrochloric acid but could not be detected and separated from the large excess of ammonia.

The amino acids were quantitatively determined by ion exchange chromatography employing the method developed by S. Moore and W. H. Stein[14].

The carboxylic acids of the first fraction were determined by means of the procedure described by W. A. Bulen *et al.*[15]. This fraction was neutralized with sodium hydroxide and evaporated. The remaining sodium salts were dissolved in dilute sulphuric acid, and admitted to a column filled with silicia gel which had been prepared with 0·5 N sulphuric acid. The carboxylic acids were eluted with mixtures of chloroform and butanol of various concentrations in the following order: 100 ml of 5 per cent, 100 ml of 15 per cent, and 100 ml of 25 per cent *n*-butanol. The eluate was collected in samples of ca. 3 ml and titrated with 1/100 N sodium hydroxide. In preliminary tests with known amounts of free acids, the positions of formic, acetic, propionic and lactic acid were determined. The same positions were obtained when the acids were admitted as sodium salts.

The solution arising from the direct photolysis experiments was only tested for amino acids by the chromatographic test as described above. The ion exchanger method was found to give negative results since it was not sufficiently sensitive to detect the extremely small amounts of amino acids.

3. EXPERIMENTAL RESULTS

Mercury-sensitized experiments

The gases were admitted to a total pressure of nearly 1 atm. During the irradiation the lamp warmed up to ca. 50°C; the partial pressure of water vapour inside the reaction vessel was approximately 100 mm Hg corresponding to this temperature. The time of irradiation was 1 week for each experiment. In all runs using mixtures of methane, ammonia and water vapour no reactions were observed.

Table 1.

No.	Gas mixture (mm Hg)	Hg	Amino acids	Carboxylic acids	Hydro-carbons
1	CH_4 (200), NH_3 (200), H_2O	+	−	−	−
2	C_2H_6 (400), NH_3 (200), H_2O	+	+	+	+
3	C_2H_6 (400), NH_3 (200), H_2O	−	+	+	+
4	C_2H_6 (600), H_2O, (NaOH)	+	−	−	+
5	C_2H_6 (600), H_2O, (NaOH)	−	−	+	−

Neither amino acids nor other compounds could be detected.

In the runs using ethane, ammonia and water vapour, a chemical conversion became visible after 6 hr of irradiation. A separate layer of a colourless liquid was collected in the *U*-tube below the condenser (see Fig. 1). The liquid consisted mainly of higher hydrocarbons which were not examined in detail. The paper chromatographic test of the water solution showed strong spots which were determined to be due to glycine and α-alanine by comparing with standard acids. In addition, weaker spots appeared which were not identified.

The quantitative analyses yielded considerable amounts of simple amino acids and fatty acids. Nearly the same results were obtained whether or not mercury was added as sensitizer. From mixtures of ethane and water vapour, however, fatty acids and liquid hydrocarbons were formed only if mercury was present. In these runs the water in the boiling flask was made alkaline in order to retain volatile acids in solution, replacing the ammonia present otherwise.

Control experiments without irradiation which were carried out with all combinations of gases did not yield any detectable reaction products.

The experiments are compiled in Table 1, including the composition of the gas mixtures and a summary of the resulting products.

The results of the quantitative analyses of amino acids and fatty acids in μ moles are given in Table 2.

Direct photolysis

With the xenon lamp a maximum of 2×10^{16} quanta/sec was available as compared to 3×10^{18} quanta/sec with the mercury lamp. If one assumes that the quantum efficiency for the formation of amino acids is the same in both series of experiments one could expect no more than a few micrograms of amino acids even under the most favourable conditions.

The experiments and analytical procedures had therefore to be carried out with carefully sterilized apparatus and controlled by repeated blank tests. In these blank tests which alternated with the full runs, either methane or ammonia in combination with water vapour was irradiated, all other conditions being kept the same.

In the full runs the gas mixture was composed of 400 mm Hg of methane and 200 mm Hg of ammonia. For blank tests the apparatus was filled with 600 mm Hg of either methane or ammonia. The time of irradiation was approximately 30 hr in each run.

The water solution resulting from the full runs was concentrated *in vacuo* and completely transferred to a few paper strips. The solution resulting from blank tests was concentrated on a single chromatogram.

Table 2. Amounts of Amino and Fatty Acids in μmoles

Experiment No.	2	3	4
Glycine	32·0	24·5	—
α-alanine	23·0	12·0	—
α-amino butyric acid	0·5	0·2	—
Total	55·5	36·7	—
Formic acid	72	82	58
Acetic acid	203	234	136
Propionic acid	17	15	9
Total	292	331	203

In the first experiments the reaction vessel was kept at room temperature. Thus the water vapour pressure inside the vessel was restricted to ca. 20 mm Hg. Under this condition no products could be detected. The formation of compounds like organic acids should depend on the presence of OH radicals arising from the photolysis of water. For this reason, the temperature was increased to 55°C in order to increase the pressure of water vapour to about 120 mm Hg.

With this change in experimental conditions the chromatograms of the full runs showed several spots. One of them was definitely determined to be glycine, since this spot was located at the same position as the standard glycine admitted for comparison. Glycine was identified on all strips independent of the developing solvent. Another spot could be attributed to α-alanine in several chromatograms. The identification of this amino acid could, however, not be accomplished since the spot was very weak and often superimposed or shifted by other spots. The chromatograms of the blank tests did not show any spots after the window cement was coated by a Teflon seal as described above.

4. DISCUSSION OF THE RESULTS

In the mercury sensitized experiments no reaction products are obtained with mixtures of methane, ammonia and water vapour. This can be understood by comparing the cross sections for quenching the Hg-resonance. The values $3{\cdot}0\times10^{-16}$ cm² of ammonia and $1{\cdot}0\times10^{-16}$ cm² of water are 50 and 20 times greater, respectively, than the value $0{\cdot}06\times10^{-16}$ cm² of methane[16]. Therefore the excited mercury atoms react nearly exclusively with the former two gases, and only H-atoms, OH- and NH_2-radicals occur as photochemical primary products. On the other hand, methane is not attacked by radicals and H-atoms up to a temperature of 300°C as confirmed by several investigations[17].

Table 3. Absorption coefficients at NTP in units of cm^{-1}

Wavelength	1296 Å	1470 Å
Methane	460	< 1
Ammonia	100	80
Water vapour	200	18·1

The quenching cross section of ethane ($0{\cdot}11\times10^{-16}$ cm²) is not much greater than that of methane. Thus the reaction

$$Hg^x + C_2H_6 \longrightarrow C_2H_5 + H + Hg$$

yields only a small contribution to the dissociation of ethane. Ethyl and methyl radicals, however, are probably formed for the most part by the reaction of ethane with H-atoms arising from the dissociation of ammonia and water vapour:

$$C_2H_6 + H \longrightarrow C_2H_5 + H_2$$
$$C_2H_5 + H \longrightarrow C_2H_6{}^x$$
$$C_2H_6{}^x + \longrightarrow 2\,CH_3{}^{[18]}.$$

This is confirmed by the non-sensitized runs using ethane, ammonia and water vapour (runs no. 3 and 5 of Table 1), which yielded nearly the same results as the sensitized runs. In non-sensitized experiments, ammonia is the only constituent which is dissociated by direct photolysis. The quantum energy available from the mercury lamp is not sufficient to dissociate ethane and water vapour directly[18, 19]. Therefore reaction products were obtained from the runs using ethane and water vapour only if mercury sensitizer was present.

Under irradiation with the xenon resonance wavelengths, the following primary reactions occur in mixtures of methane, ammonia and water vapour:

$$CH_4 + h\nu \longrightarrow CH_3 + H$$
$$NH_3 + h\nu \longrightarrow NH_2 + H$$
$$H_2O + h\nu \longrightarrow OH + H.$$

The absorption coefficients at the wavelengths 1296 and 1470 Å taken from the measurements of Watanabe *et al.*[20] are listed in Table 3.

The ratio of the light intensities at the wavelengths 1470 and 1296 Å can be estimated to be in the order of 10:1. Considering the values of absorption coefficients it is reasonable to

assume that the radicals CH_3, NH_2 and OH are primarily formed at a comparable rate from the gas mixtures used in the runs.

As a final product, only glycine could be definitely determined, but even this in extremely small amounts compared with the sensitized experiments. This may be explained mainly by the small quantum yield available from the xenon lamp.

In addition, C_2H_5-radicals are formed primarily in experiments with ethane, in contrast with CH_3-radicals in the experiments with methane. Assuming that in the former case, the C–C bond of the glycine molecule is readily provided by the C_2H_5-radicals*), this C–C bond must first appear in the case of reactions with methane. For this purpose at least one additional step would be necessary. The formation of glycine would therefore be more probable if C_2H_5-radicals are primarily present rather than CH_3-radicals.

The formation of carboxylic acids in the sensitized experiments using only ethane and water vapour permits the conclusion that the COOH group can be built up by a mechanism which does not contain intermediate nitrogen compounds.

Acknowledgements—These investigations were supported by the Geophysics Research Directorate of the US Air Force under Contract No. AF61(514)-1079 with the European Office, and by the Deutsche Forschungsgemeinschaft.

We would like to thank Dr. Krampitz[21] of the Institut für Anatomie und Physiologie der Haustiere, University of Bonn, for carrying out the quantitative analysis of the amino acids.

REFERENCES

1. A. I. Oparin, *The Origin of Life*, translated by S. Morgulis (2nd Ed.). Dover Publications, New York (1953).
2. H. C. Urey, *The Planets*. Oxford University Press (1952).
   H. C. Urey, *Proc. Nat. Acad. Sci., Wash.* **38**, 351 (1952).
3. W. Kuhn, *Chem. Ber.* **89**, 303 (1956).
   W. Kuhn and A. Rittmann, *Geol. Rdsch.* **32**, 215 (1941).
4. J. D. Bernal, *The Physical Basis of Life*. Routledge, London (1957).
5. M. Calvin, *Naturwissenschaften* **43**, 387 (1956).
6. S. L. Miller, *J. Amer. Chem. Soc.* **77**, 2351 (1955).
7. K. Heyns, W. Walter and E. Meyer, *Naturwissenschaften* **44**, 385 (1957).
8. H. A. Ratsack, *Z. phys. Chem. Lpz.* **206**, 285 (1957).
9. W. Groth and H. v. Weyssenhoff, *Naturwissenschaften* **44**, 510 (1957).
10. W. Groth, H. Ihle and A. Murrenhoff, *Angew. Chem.* **68**, 605 (1956).
11. R. N. Smith, P. A. Leighton and W. G. Leighton, *J. Amer. Chem. Soc.* **61**, 2299 (1939).
    L. B. Thomas, *J. Amer. Chem. Soc.* **62**, 1879 (1940).
    L. Küchler and H. Pick, *Z. phys. Chem.* **B45**, 116 (1940).
12. P. Harteck and R. Oppenheimer, *Z. phys. Chem.* **B16**, 77 (1932).
13. W. Groth, *Z. Elektrochem.* **42**, 533 (1936).
    W. Groth, *Z. phys. Chem.* **B37**, 307 (1937).
    W. Groth and W. Scharfe, *Zs. phys. Chem. Frankf.* **2**, 142 (1954).
14. S. Moore and W. H. Stein, *J. biol. Chem.* **211**, 893 (1954).
15. W. A. Bulen, J. E. Varner and R. C. Burrel, *Analyt. Chem.* **24**, 187 (1952).
16. J. R. Bates, *J. Amer. Chem. Soc.* **52**, 3852 (1930).
    M. G. Evans, *J. Chem. Phys.* **2**, 445 (1934).
17. P. Harteck, K. F. Bonhoeffer and K. H. Geib, *Z. phys. Chem.* **A139**, 64 (1928) and **A170**, 1 (1934).
18. B. Darwent, *J. Chem. Phys.* **18**, 1532 (1950).
    Phys. **16**, 381 (1948).
    B. Darwent, *J. Chem. Phys.* **18**, 1532 (1950).
19. H. Senftleben and J. Rehren, *Z. phys. Chem.* **37**, 529 (1926).
20. K. Watanabe, M. Zelikoff and E. C. Y. Inn, *Air Force Cambr. Res. Center, Techn. Rep.* 53/23 (1953).
21. G. Krampitz and R. Müller, *Z. Tierernähr.* **14**, 114 (1959).

* This assumption is supported by the fact that the analysis of the carboxylic acids yielded an amount of acetic acid 3 times higher than the amount of formic acid.

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 86–98. Printed in Great Britain.

IONIZATION OF A STATIC INTERPLANETARY GAS AND EXPECTED EMISSION LINES FROM THIS GAS

H. N. RUNDLE

The Observatories, Cambridge University, England

(*Received* 3 *August* 1959)

Abstract—The degree of ionization of hydrogen and helium in a static interplanetary gas is calculated and the expected intensities of the more intense helium emission lines in the visible and infra-red regions are given as a function of electron density and kinetic temperature. A zodiacal light spectrum recently obtained places an upper limit on the static electron density for given kinetic temperatures. Future high resolution observations of λ 5876 or λ 10,829 Å could give useful upper limits on the static electron densities for kinetic temperatures below 10^{5}°K. Comments are made on the observations of H_2.

1. INTRODUCTION

Very little is known about the interplanetary gas. It is generally believed that fast solar streams or corpuscular radiations are responsible for intense terrestrial magnetic storms and at least part of auroral activity. In addition to these fast streams, an interplanetary gas has been suggested mainly from observations of the zodiacal light. Whipple and Gossner[1] suggested in 1949 that the electrons resulting from the ionization of this gas are responsible for most of the polarized component of the zodiacal light and that the light scattered by interplanetary dust is only weakly polarized. Whistler observations have given electron densities near the earth and observations of radio star scintillations and the solar corona have given information on electron densities near the sun. The electron density of the interplanetary gas is thought to be about 100 to 1000 cm^{-3} at the earth's orbit. It is often believed that the degree of ionization is high and that the kinetic temperature is between 5000 and 10^{6}°K, although in 1959 Kupperian *et al.*[7] suggested a temperature of about 400°K.

It is not known whether the gas is static or moving radially with respect to the sun. As there have been no reported fluctuations in the polarization of the zodiacal light, it is usually assumed that the gas is static. On the other hand, comet tail accelerations have been explained by a gas streaming away from the sun, and the ionization and brightness of comet nuclei have also been explained by the impact of the particles of such a streaming gas. A summary of the above known facts and views concerning an interplanetary gas has been given by Piggott and Allen[2] in 1957.

The physical properties of this gas are of fundamental importance for the study of any phenomena involving solar–terrestrial relationships and these properties could be determined if we could observe an emission line produced by the gas. Smyth[3] attempted in 1954 to detect an extra absorption line in the solar spectrum which could be attributed to corpuscular radiation but this difficult observation did not give any positive results. There has been no attempt to observe emission lines from a solar stream although Jones[4] has observed the *H* and *K* emission lines of Ca II in the night sky and these could be connected with such a stream. This communication discusses the possibility of observing a useful emission line from a static interplanetary gas.

The possibilities of observing an emission line can be assessed only if the degree of ionization is known for the particular element under dis-

cussion. Since hydrogen and helium are the two most abundant elements, they are obvious choices for investigation. Of these two elements, helium is the more interesting because of the triplet states which give rise to emission lines in the visible and infra-red regions. Hence, it is instructive to investigate the possibility of observing a line arising from these triplet states. On the other hand, the majority of singlet atoms are in the ground state and hence the strongest singlet emission line is in the far ultra-violet. The intensity of other singlet fluorescent lines depend on the intensity of the exciting radiation around 550 Å. Similarly, for hydrogen the strongest emission line is Lyman α and the intensity of the Balmer lines depend on the intensity of the exciting Lyman radiation. The first step in this investigation is the calculation of the degree of ionization of hydrogen and helium.

2. GENERAL REMARKS ON THE IONIZATION OF HELIUM AND HYDROGEN

There have been three estimates of the ionization of interplanetary hydrogen. Opik[5] concluded in 1954 that hydrogen is almost completely ionized out to approximately the orbit of Jupiter. He used an electron kinetic temperature of about 1000°K obtained by considering the distribution of interplanetary matter perpendicular to the ecliptic. He assumed an electron density of 600 cm^{-3} and a flux of photons of $5 \cdot 10^{11}$ cm^{-2} sec^{-1} with energies greater than 13·5 eV at a distance of one A.U. from the sun. Elsässer[6] in his review of 1958 on interplanetary material, reaches a similar conclusion concerning the degree of ionization. He assumed an electron kinetic temperature of 7000°K and black body radiation from the sun at the same temperature. Both calculations neglect the variation of the atomic absorption coefficient with frequency and the ionization by collision with electrons. Kupperian *et al.*[7] in 1959 arrived at a similar conclusion for their cold gas. Several investigators have calculated the ionization and population of excited states of hydrogen and helium for solar conditions[8–11], and for gaseous nebular conditions[12]. The conditions in interplanetary space are intermediate between those existing in the sun and in gaseous nebulae, because collision excitations and ionization can be important and the radiation is moderately diluted.

The temperature, electron densities and dilution of radiation are kept as variables and the following assumptions are made in the calculation of the ionizations. As hydrogen has the highest cosmic abundance and is almost completely ionized it is assumed that the electron density is equal to the total hydrogen density. This means that the gas consists mainly of electrons and protons. The interaction of dust with the gas is neglected because the dust density is about 10^{-15} particles cm^{-3} and the electron density is about 100 cm^{-3}. The gas is considered to be optically thin. The ionization processes considered are photo-ionization and ionization by electron collision. The cross-sections for ionization by electron and proton collision are nearly the same, as was shown by Dalgarno and McDowell[13] in 1955, and thus the ratio of the rates of ionization by collisions of electrons and protons is equal to (proton mass/electron mass)$^{1/2}$. Hence ionization by proton collision is neglected. The recombination process is considered to be photo-recombination. Three body collisions and charge transfers are neglected because of the small densities and also because most of the gas is in the form of electrons and protons. The electrons are assumed to have a Maxwellian velocity distribution although this is not necessarily true. Stimulated emission is neglected.

In the calculations the ionization and recombination rates and, if necessary, the excitation rates are considered and solutions are found for the steady state. Therefore, the results are true only if the gas exists for at least the life of a free atom or an ion, which may be measured in years or centuries, i.e. the gas must be static.

3. THE FORMULAE AND DATA EMPLOYED

Let the number of s-times ionized atoms in the mth quantum state be $n_{s,m}$. Then it is well

known[14, 15] that the rate of photo ionization/cm³/sec is

$$Z_{pi}=\frac{n_{s,m}}{h\,r^2}\int_{\nu_m}^{\infty} I_\nu W_e\,\frac{(\mathscr{K}_\nu)_m}{\nu}\,d\nu$$

where $I_\nu W_e/r^2$ is the incident radiation in the spectral range of ν to $\nu+d\nu$, r is the distance from the sun in A.U. and W_e is the dilution factor for $r=1$. ν_m is the frequency at the ionization limit and $(\mathscr{K}_\nu)_m$ is the atomic absorption coefficient for the quantum state m. Now $(\mathscr{K}_\nu)_m=\mathscr{K}_m(\nu_m/\nu)^p$ where $p\leqslant 3$ thus

$$Z_{pi}=\frac{n_{s,m}}{r^2}\,\frac{\mathscr{K}_m}{hc\,(\lambda_m)^p}\int_0^{\lambda_m} I_\lambda W_e\lambda^{p+1}d\lambda \qquad (1)$$

The electron collision ionization rate is given by $Z_{ci}=n_{s,m}\,q_{s,m}\,\nu dn_e$ where $q_{s,m}$ is the cross-section and dn_e is the number of electrons with velocities between ν and $\nu+d\nu$. Now, for hydrogen and helium[16]

$$q_{s,m}=4b\,(\pi a_0^2)\,\frac{\varepsilon_e-\chi_{s,m}}{\chi_{s,m}\varepsilon_e^{\,2}}$$

where πa_0^2 is the area of the first Bohr orbit, ε_e and $\chi_{s,m}$ denote the incident electron energy and the ionization potential, respectively, in rydbergs and b is a constant, taken to be $\frac{1}{2}$ following Woolley and Allen[17]. Upon making the required substitutions and integrations with the assumption of a Maxwellian velocity distribution for the electrons

$$Z_{ci}=2{\cdot}34\,.\,10^{-4}\,\frac{n_{s,m}\,n_e}{\chi_{s,m}\sqrt{T_e}}\,E\left(\frac{\chi_{s,m}}{k\,T_e}\right) \qquad (2)$$

where n_e is the electron density, T_e is the electron kinetic temperature, k is Boltzmann's constant and $E_{(2)}=e^{-x}+xE_1(-x)$ where

$$-E_1(-x)=\int_x^{\infty}\frac{e^{-t}}{t}\,dt.$$

The photo recombination rate per cm³ to the mth quantum state of an s times ionized atom is $z_{R,n}=n_{s+1}\int(\beta_\nu)_m\,\nu dn_e$ where $(\beta_\nu)_m$ is the cross-section and n_{s+1} is the density of $(s+1)$ times ionized atoms. Now from Ambartsumyan[14] or Unsöld[15]

$$(\beta_\nu)_m=\frac{g_{s,m}\,\nu^2\,h^2}{u_{s+1}\,c^2\,m_e^2\,v^2}\,(\mathscr{K}_\nu)_m$$

where $g_{s,m}$ is the statistical weight of the mth quantum state of an s times ionized atom, u_{s+1} is the partition function of the $(s+1)$ times ionized atom and the other symbols have already been defined or have their conventional designations. Upon making the required substitutions $Z_{R,n}$ is obtained for different p.

$$[Z_{R,n}]_{p=2}=1{\cdot}08\,.\,10^{-4}\,\frac{g_{s,m}}{u_{s+1}}\,n_e\,n_{s+1}\,\frac{\mathscr{K}_m}{\lambda_m^2\,T_e^{\frac{1}{2}}} \qquad (3)$$

and

$$[Z_{R,n}]_{p=3}=1{\cdot}55\,.\,10^{-4}\,\frac{g_{s,m}}{u_{s+1}}\,.\,n_e\,n_{s+1}\,\frac{\mathscr{K}_m}{\lambda_m^3\,T_e^{3/2}}\times$$

$$\times\exp\left(\frac{\chi_{s,m}}{kT_e}\right)\left[-E_1\left(-\frac{\chi_{s,m}}{kT_e}\right)\right] \qquad (4)$$

The case for $p=1$ is not required. For a hydrogen like atom, $p=3$, $u_{s+1}=1$; $g_{s,m}=2n^2$. The absorption coefficient $\mathscr{K}_n$ is given by

$$\mathscr{K}_n=\frac{32\pi^2e^6RZ^4}{3\sqrt{3}ch^3}\,g'\,\frac{1}{\nu^3n^5}$$

where g' is the Gaunt factor of about unity[14, 15]. Ze is the net nuclear charge and R is the rydberg frequency. From (4) the photo recombination rate for hydrogen is

$$[Z_{R,n}]_H=3{\cdot}22\times10^{-6}\,n_en_{s+1}\,Z^4\,M_n(n,T) \qquad (5)$$

where $M_n(n,T)=\dfrac{\exp\left(\frac{\chi_n}{kT_e}\right)\left[-E_1\left(-\frac{\chi_n}{kT_e}\right)\right]}{n^5\,T_e^{3/2}}$

The function $M_n(n,T)$ has been tabulated for a few of the desired temperatures by Cillié[18] who in 1932 investigated hydrogen equilibrium in gaseous nebulae.

The rates of excitation by collisions with electrons are also required. Various cross-sections are given by Allen[16]. The cross-sections for optically forbidden excitations are given by $Q_{A\to D}=0{\cdot}237b\,g_D\,\pi a_0^2\,\dfrac{\varepsilon_e-\chi_{AB}}{\varepsilon_e^{\,2}}$

For optically allowed excitations the cross-sections are $Q_{A\to B}=\dfrac{12{\cdot}0fb\pi a_0^2}{\chi_{AB}}\dfrac{\varepsilon_e-\chi_{AB}}{\varepsilon_e^3}$

where g_B and f are the multiplicity of the upper state and the absorption oscillator strength, respectively. b is a factor determined experimentally and is given by Allen[16]. The other symbols are defined above. Now, the excitation rates are

$$Z_{A\to B}=n_A\int Q_{A\to B}\,v\,dn_e$$

or

$$Z_{A\to B}=\frac{2{\cdot}04\,.\,10^{-6}\,n_e n_A b g_B}{\sqrt{T_e}}\,E\left(\frac{\chi_{AB}}{kT_e}\right) \qquad (6)$$

for optically forbidden excitations. For optically allowed excitations

$$Z_{A\to B}=\frac{1{\cdot}41\,.\,10^{-3}\,n_e n_A f b}{\chi_{AB}\sqrt{T_e}}\,E\left(\frac{\chi_{AB}}{kT_e}\right); \qquad (7)$$

χ_{AB} is in electron volts. The de-excitation and excitation rates follow a Boltzmann law such that

$$\frac{Z_{A\to B}}{n_A g_B}=\frac{Z_{B\to A}}{n_B g_A}\exp-\left(\frac{\chi_{AB}}{kT_e}\right)$$

The data employed are given in Table 1. $\mathscr{H}_m$ and p were obtained from Goldberg[19] and Allen[16]. The value of $\int I_\lambda W_e\lambda^{p+1}d\lambda$ from 0 to λ_m was estimated from the theoretical work described in 1955 by de Jager[20] and the rocket measurements of Friedman[21] which are in reasonable agreement with each other. Since the solar radiation in the above integral is weighted with λ^{p+1} the solar radiation just below the ionization limit λ_m is the most important and the fluctuations of solar radiation at shorter wavelengths will have little effect on the value of the integral. For the conditions of interplanetary space, the most important integrals are those associated with the ground states of H I, He I and He II. Unfortunately the intensity of the solar radiation in the region of λ_m for these particles can be measured only from rockets or satellites and is not yet well known.

4. HYDROGEN IONIZATION

It was estimated from Friedman's[21] measurements of Lyman α that a negligible number of atoms are present in the $n=2$ level of hydrogen. Thus it is safe to conclude that ionization from the ground state only need be considered. Recombinations, however, occur to all quantum levels and the recombined atoms cascade to the ground level. Of the negligibly few atoms excited from the ground state, most will fall back to the ground state instead of being excited again or becoming ionized, owing to the long time between electron collisions and the diluted radiation. Upon solving the equilibrium equation, $Z_{pi}|_{n=1}+Z_{ci}|_{n=1}=\sum_{n=1}^{\infty}Z_{R,n}$, and using equations (1), (2) and (5) the ionizations of Fig. 1 were found. $Z_{R,n}$ was neglected for $n>10$ and this neglect introduces a small error in $\sum_{n=1}^{\infty}Z_{R,n}$ as shown by Cillié[18].

Table 1. Data Employed for Ionization Calculations.

Atom and state	$\mathscr{H}_m$ (cm²)	p	$\chi_{e,m}$ (eV)	λ_m (Å)	$\int_0^{\lambda_m} I_\lambda W_e \lambda^{p+1} d\lambda$ (erg cm^{p-1} sec^{-1})
H I 1	$6{\cdot}3\times10^{-18}$	3	13·6	912	$1{\cdot}2\times10^{-23}$
He I 1^1S	$7{\cdot}6\times10^{-18}$		24·45	506	$2{\cdot}0\times10^{-18}$
2^3S	$2{\cdot}8\times10^{-18}$		4·75	2670	$2{\cdot}3\times10^{-6}$
2^3P	$16{\cdot}0\times10^{-18}$	2	3·61	3440	$1{\cdot}65\times10^{-9}$
2^1S	$10{\cdot}5\times10^{-18}$		3·95	3140	$4{\cdot}6\times10^{-10}$
2^1P	$13{\cdot}0\times10^{-18}$		3·5	—	—
He II 1^1S	$1{\cdot}8\times10^{-18}$	3	54·4	230	$8{\cdot}0\times10^{-25}$

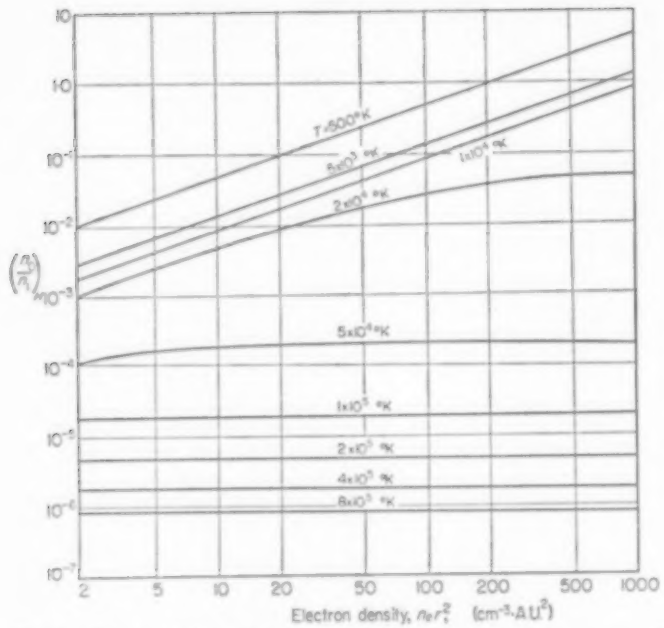


Fig. 1. The ionization of hydrogen in a static interplanetary gas. T refers to the kinetic temperature T_e of the gas; n_0 and n_1 are the densities of neutral and ionized hydrogen, respectively.

It should be noted that collision and photo-ionizations are of approximately equal importance at a temperature of about 2×10^4 °K. Table 2 gives the lifetime of a free hydrogen atom and of a proton as a function of temperature and electron density.

5. HELIUM I EXCITATIONS AND POPULATION OF STATES

The ionization of neutral helium is complicated by the presence of the singlet and triplet levels and the metastable 2^1S level shown in Fig. 2. Before determining the ionization we must examine the relative populations of the different levels. The two metastable levels and the ground state have populations which are decided by electron collisions, ionization, recombination and photo-excitations. The number of atoms entering a level above the metastable levels from recombination and collisional excitations is negligible compared to the numbers entering the level by photo-

Table 2. Lifetime of Hydrogen Atoms and Protons in a Static Interplanetary Gas.

T_e (°K)	5×10^3	5×10^3	1×10^4	2×10^4	5×10^4	1×10^5	2×10^5	4×10^5	8×10^5
Life of a free H I atom ($1/n_0 n_e$ years)	—— $63{\cdot}5\, n_e r^2$ ——				46	9·0	4·1	2·7	2·7
Life of a proton ($10^5/n_e$ years)	0·13	0·47	0·75	1·2	2·4	4·3	8·2	16	33·0

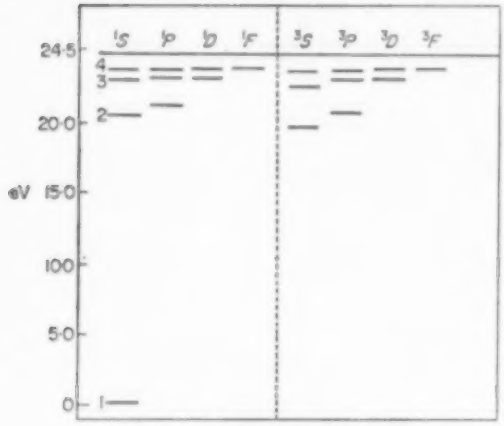


Fig. 2. The He I energy levels.

excitation. Hence the levels above the two metastable levels will have populations essentially determined by the incident radiation and this population will follow a Boltzmann distribution to a good approximation. If the mean temperature of the sun is assumed to be 6000°K one can conclude from the Boltzmann distribution that $\frac{n_{2^3p}}{n_{2^3s}} = 1{\cdot}8 \times 10^{-6}/r^2$, etc. Similar ratios are found for the other triplet and singlet levels.

The excitation and de-excitation rates between metastable and ground state are required. The cross-sections for transitions from the 1^1S level to the singlet and triplet levels are known for many values of n[16]. Transitions from the 1^1S level to levels above $n=2$ are neglected because their cross-sections are small compared with those for the transitions to the $n=2$ levels. Now the atoms entering the 2^1P level will fall to the 2^1S and 1^1S levels in the ratio of the transition probabilities, which is 0·001–1, respectively. Thus atoms will be easily drained off the 2^1S level to the 1^1S level via the 2^1P level but it is more difficult for them to return. The rate of the transition $1^1S \rightarrow 2^1S$ including the transition $1^1S \rightarrow 2^1P \rightarrow 2^1S$ is α_1 electron^{-1} (initial atom)$^{-1}$. It was calculated neglecting photo-excitation of the 1^1S level to the 2^1P level, because the intensity of the exciting radiation of 584° Å is variable and cannot be measured from the earth's surface. Thus the calculated population of the excited singlet levels may be too low for the lower kinetic temperatures when photo-excitation becomes more important than collisional excitation. Fortunately, α_1 is important for the singlet level populations only, as is shown later. The transition $2^1S \rightarrow 1^1S$ occurs chiefly via the 2^1P level, and its rate, α_2 is $10/r^2$ sec^{-1} (initial atom)$^{-1}$ for all kinetic temperatures. β_1 electron^{-1} (initial atom)$^{-1}$ is the rate of the transition $1^1S \rightarrow 2^3S$ including the transition $1^1S \rightarrow 2^3P \rightarrow 2^3S$. For the inverse transition the 2^3P level is neglected due to its small population relative to the 2^3S level and β_2 (initial atom)$^{-1}$ electron^{-1} is the rate of the $2^3S \rightarrow 1^1S$ transition. Fig. 3 shows α_1, β_1 and β_2 as a function of the temperature. Equations (6) and (7) were used throughout the calculations.

The cross-sections of other singlet-triplet transitions are unknown, but it was felt that the transitions from the 2^1S state to the triplet state and from the 2^3S state to the singlet state should be accounted for. The $2^1S \rightleftharpoons 2^3S$ transitions were considered only as an indication of the effect of these transitions. The cross-section was assumed to be the same as for the 1^1S state. The energy dependence is probably similar and the experimentally determined b is probably about unity since the case of $b \ll 1$ applies when $\Delta n > 1$, and $b > 1$ violates the limiting cross-section rule. Fortunately, the rates of the excitation and de-excitation transitions, γ_1 and γ_2, can be neglected in the solution of the ionization equation for the conditions met in interplanetary space. Radiative transitions between singlet and triplet levels can be neglected. The most important such line is at 591 Å and is due to the transition of $2^3P \rightarrow 1^1S$; it was observed by Hopfield[22] in 1930 in the laboratory.

6. IONIZATION AND RECOMBINATION OF HELIUM I

Because of the dilute radiation, the majority of the atoms of helium I are in the 1^1S, 2^1S and 2^3S levels and ionization from these levels only need be considered. The photo- and collisional-

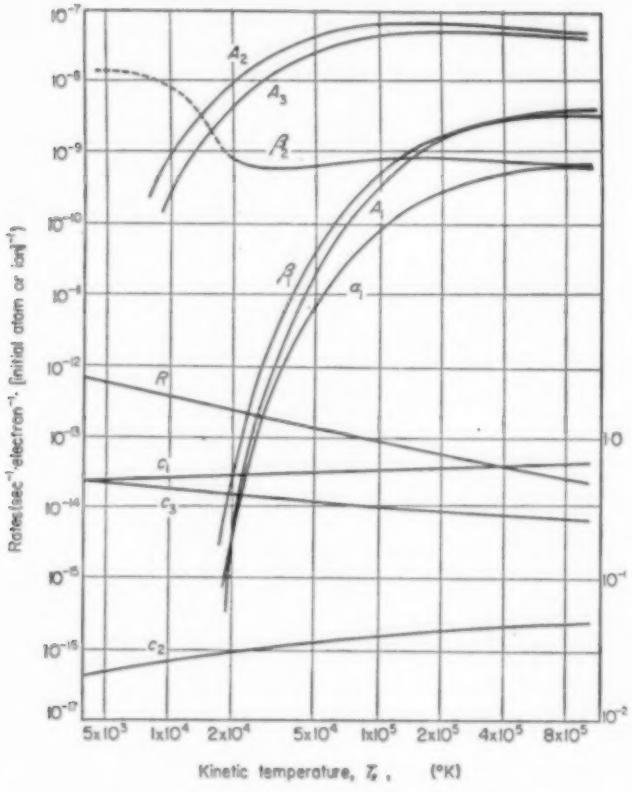


Fig. 3. The rates of collisional excitations, de-excitation and ionizations and the rate of photo recombinations. The quantities C_1, C_2 and C_3 are also shown, their abscissa being on the left.

ionization rates were calculated from equations (1) and (2). The photo-ionization rate is denoted by P/r (initial atom)$^{-1}$ and the collisional-ionization rate is denoted by A electron^{-1} (initial atom)$^{-1}$ and subscripts 1, 2 and 3 refer to the levels 1^1S, 2^1S and 2^3S, respectively. We obtain $P_1 = 3 \times 10^{-9}$ $P_2 = 2{\cdot}5 \times 10^{-2}$ and $P_3 = 1{\cdot}2 \times 10^{-3}$ sec^{-1} (initial atom)$^{-1}$ and the quantities A_1, A_2 and A_3 are given in Fig. 3.

The rate of recombination denoted by R electron^{-1} (initial atom)$^{-1}$ gives the recombinations to all levels up to $n = 10$. For the quantum levels with $n > 2$, the hydrogen approximation is satisfactory[15]. For $n = 2$ and 1 equation (3) must be used. These recombining atoms will eventually find their way to one of the metastable or ground levels. The fractions ending up in the 1^1S, 2^1S and 2^3S states are denoted by C_1, C_2 and C_3, respectively where $C_1 + C_2 + C_3 = 1$. The calculation of these coefficients assumes that the number of atoms recombining to any fine structure level is proportional to the statistical weight of that level for a given total quantum number. Thus $C_3 = 3/4$ (atoms recombining to the $n = 2, 3 \ldots$ levels/total number of recombining atoms), and C_1 and C_2 depend on the relative values of the transition probabilities of the $^1P \rightarrow 2^1S$ and $^1P \rightarrow 1^1S$ transitions. R, C_1, C_2 and C_3 are shown in Fig. 3.

7. THE IONIZATION EQUATIONS AND THEIR SOLUTIONS

The ionization equations state that in equilibrium the rate at which atoms leave each of the metastable and ground states is equal to the rate of re-entry into these states. The equations are

$$\begin{pmatrix} \alpha_1+\beta_1+S_1 & -\dfrac{\alpha_2}{n_e r^2} & -\beta_2 \\ -\alpha_1 & \dfrac{\alpha_2}{n_e r^2}+\beta_2+S_2 & -\gamma_1 \\ -\beta_1 & -\gamma_1 & \gamma_1+\beta_2+S_2 \end{pmatrix} \times \times \begin{pmatrix} n_{1^1S} \\ n_{2^1S} \\ n_{2^3S} \end{pmatrix} = n_1 R \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} \quad (8)$$

where n_1 is the density of helium II and $S=(P/n_e r^2)+A$. For $n_e r^2 \leqslant 10^3$, the solutions are

$$\begin{aligned} \frac{n_{1^1S}}{n_1} &= \frac{(C_1+C_2)\,R n_e r^2}{P_1+n_e r^2\,(\beta_1+A_1)} \\ \frac{n_{1^1S}}{n_1} &= \left(\frac{n_{2^1S}}{n_1}\,\alpha_1+C_2 R\right)\frac{n_e r^2}{\alpha_2} \\ \frac{n_{2^3S}}{n_1} &= \frac{C_3 P_1+n_e r^2\,(\beta_1+C_3 A_1)}{P_1+n_e r^2\,(\beta_1+A_1)} \end{aligned} \quad (9)$$

Fig. 4 gives numerical values of these ratios as a function of temperature and $n_e r^2$.

The ionization of He II is completely analogous to the hydrogen case and the ratios of n_1/n_2 are shown in Fig. 4 where n_2 is the density of helium III.

Neglecting the He I particles, the lifetime of a He II or He III particle is of the same order of magnitude as for the hydrogen case shown in Table 2. The life time of a He I particle in ground state is $10r^2/n_0$ years for T_e less than 2×10^4 °K and decreases to $30/n_0 n_e$ years for T_e above 10^5 °K and high electron densities. The life time of a He II particle due to recombinations to He I is of the order of $10^5/n_1 n_e$ years. Hence total equilibrium will occur after the gas has existed for about $10^5/n_1 n_e$ years.

8. EMISSION LINES FOR HYDROGEN AND HELIUM IN A STATIC INTERPLANETARY GAS

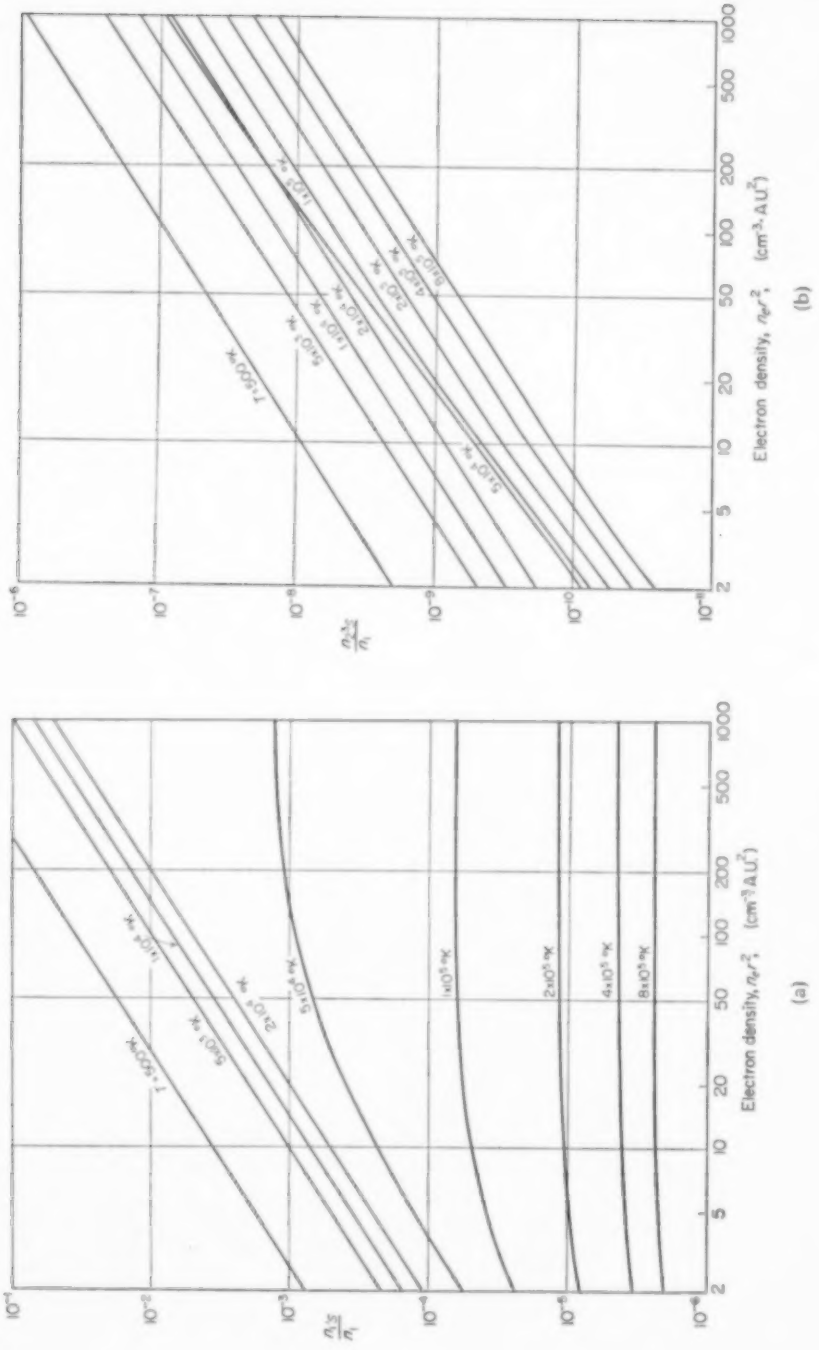
First, a few comments on the hydrogen ionization are instructive. Kupperian *et al.*[23] suggested in 1955 that $n_{\text{II}}>1$ atom cm^{-3}. In 1959[7] they suggested $n_{\text{II}}>0{\cdot}2$ atoms cm^{-3}. Other workers[24, 25] obtain similar results. Therefore we will assume that $n_{\text{II}}>0{\cdot}1$ atoms cm^{-3}. If $n_e<10^3$ cm^{-3}, then $n_{\text{II}}/n_1>10^{-4}$ and, from Fig. 1, $T_e<5\times10^4$ °K.

Recombining atoms give rise to less than 5×10^5 quanta sec^{-1} cm^{-2} for any helium I emission line. This intensity is 0·5 rayleighs* which can be neglected. Thus excitations of hydrogen and helium atoms will determine any emission line that could be observed.

For hydrogen, the intensity of any emission line will depend on the intensity of the exciting Lyman radiation. In 1958 Kupperian *et al.*[23, 27] measured Lyman α in the night sky and it appears as though 10^{-3} ergs sec^{-1} cm^{-2} might arise from interplanetary space†. This corresponds to 500 rayleighs. The strongest emission line in the visible range will be Hα and Shklovsky[25] has calculated its intensity to be 20 rayleighs from a cold interplanetary gas. This emission would make the Hα absorption line of the zodiacal light shallower than the solar Hα line. This emission might also be seen in night sky spectra when the solar Lyman emission is strong. Elvey[28] in 1950 identified a line in the night sky spectrum as Hα and Meinel[29] pointed out in the same year that this line coincided with an intense OH emission. It is claimed that Hα has an intensity between 5 to 20 rayleighs in the night sky at Zvenigorod as referred to by Kupperian *et al.*[7]. Shklovsky[30] gave reasons in 1958 for expecting Hα emission from the upper atmosphere during twilight or in the night sky as well as from

* One rayleigh (symbol R) is equivalent to 10^6 quanta being emitted in all directions from a 1 cm^2 column in the line of sight[26].

† This reasoning is invalid if there is a concentration of interplanetary matter around the earth or a second layer of hydrogen in the terrestrial atmosphere above 200 km.


Electron density, $n_e r^2$; (cm^{-3}-A.U.^{2})
(b)
$\frac{\Delta\nu}{\nu}$
$T=500°K$
$\chi_e=2\times10^{-4}$
$\chi_e=1\times10^{-4}$
$\chi_e=5\times10^{-5}$
$\chi_e=2\times10^{-5}$
$\chi_e=1\times10^{-5}$
$\chi_e=5\times10^{-6}$
$\chi_e=2\times10^{-6}$
$\chi_e=1\times10^{-6}$
Electron density, $n_e r^2$; (cm^{-3}-A.U.^{2})
(a)
$\frac{\Delta\nu}{\nu}$
$T=500°K$
$\chi_e=2\times10^{-4}$
$\chi_e=1\times10^{-4}$
$\chi_e=5\times10^{-5}$
$\chi_e=2\times10^{-5}$
$\chi_e=1\times10^{-5}$
$\chi_e=5\times10^{-6}$
$\chi_e=2\times10^{-6}$
$\chi_e=1\times10^{-6}$
$\chi_e=5\times10^{-7}$
$\chi_e=2\times10^{-7}$
$\chi_e=1\times10^{-7}$

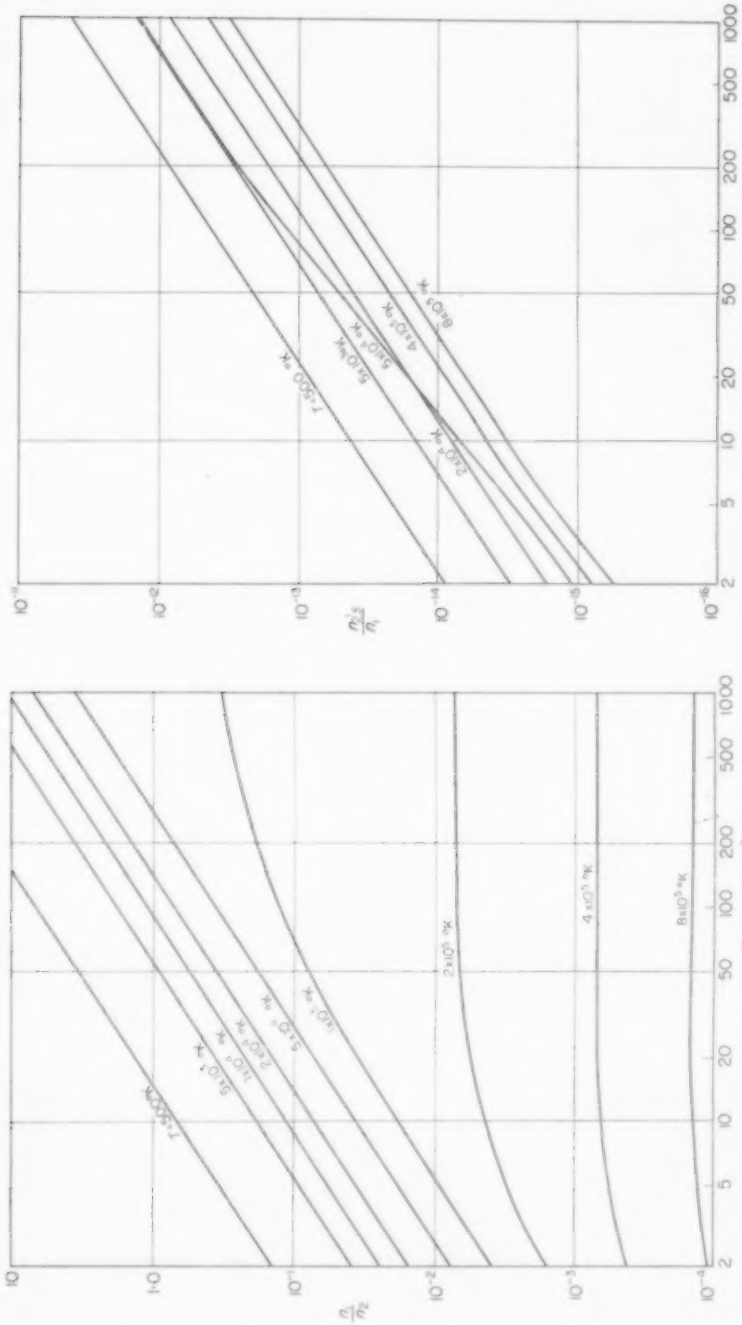


Fig. 4. The ionization of helium in a static inter-planetary gas. T refers to the kinetic temperature T_e of the gas; n_0, n_1 and n_2 are the densities of neutral, singly ionized and doubly ionized helium, respectively.

interplanetary gas. High resolution is required to separate any H_2 emission from the strong OH bands and its interpretation is difficult. However, if its bandwidth could be measured a temperature would be obtained.

The singlet levels of helium I have populations several orders of magnitude less than the populations of the triplet levels. Hence, transitions among the singlet levels will be relatively weak, except for the resonance lines between the P levels and the ground state which are near 500 Å. However, if the solar helium emission lines increase in intensity by a few orders of magnitude, the $2^1P \longrightarrow 2^1S$ transition at 2·06 μ might be seen for high electron densities and low temperatures. The four strongest emission lines from triplet helium are the $3^3S \longrightarrow 2^3P$, $3^3P \longrightarrow 2^3S$, $3^3D \longrightarrow 2^3P$, and $2^3P \longrightarrow 2^3S$ transitions at 7065, 3889, 5876 and 10,829Å, respectively. Neglecting fluorescence, the intensity of these lines can be calculated from the previously mentioned Boltzmann population distribution of the triplet levels and the ionization of helium I and II, which is a function of $n_e r^2$ and the kinetic temperature of the electrons. It is assumed that the electron density is essentially the same as the net hydrogen density and that helium and hydrogen occur in the cosmic abundance of 1 to 10 in the number of atoms. The path length is taken to be 10^{13} cm or 1 A.U. With these assumptions, the following equation (10) gives the intensity of λ 10,829 in rayleighs where the densities used refer to helium

$$I_{\lambda\ 10829} = \frac{3{\cdot}6 \times 10^8}{r^2} \frac{\dfrac{n_{2^3S}}{n_1}}{1 + \dfrac{n_{2^1S}}{n_1} + \dfrac{n_2}{n_1}} n_e \text{ rayleighs} \quad (10)$$

An average value $\overline{n_e r^2}$ is used for the value of the argument of the density ratios. Fig. 5 shows the expected intensity of the λ 10,829 Å line as a function of $\overline{n_e r^2}$ and T_e. The intensities of the λλ 5876, 7065 and 3889 Å lines are 0·19, 0·017 and 0·016 times the intensity of the λ 10,829 line, respectively.

It should be emphasized that the accuracy of the calculated ionization and intensities depend upon the accuracy of the assumed solar radiation near the ionization limits of H I, He I and He II. For temperatures above 10^5 °K ionization by collision is more important than photo ionization; thus errors in the intensity of the solar radiation will be important below this temperature only. The effect of an over estimate or under estimate of this radiation is to raise or lower the effective kinetic temperature, respectively, and to alter the temperature where photo and collisional ionizations are of equal importance.

9. OBSERVATION OF THE EMISSION LINES

The brightness of the zodiacal light continuum determines the weakest emission line capable of observation. The Doppler widths of the lines are approximately 1/4 to 1 Å. The brightness of the zodiacal light continuum is 4 to 5 rayleighs per Å at λ 6300 Å for an elongation of about 35° if Blackwell's figures[31] are used. Then the weakest observable line probably has an intensity of 1/4 to 1 rayleigh at this elongation. This fact limits observable emissions to cases where there are high electron densities or low temperatures. Observations from the ground are made more difficult yet by the presence of intense OH bands in the night air glow, although high resolution techniques can overcome this trouble.

In recent zodiacal light spectra obtained by Blackwell and Ingham (unpublished) there is no indication of λ 5876 emission and hence its intensity cannot be greater than about 1 rayleigh. The elongation of the line of sight was about 30°, hence r varies from unity to 0·5 over the path length of 10^{13} cm. Therefore, $1/r^4$ varies from unity to 16 and $\overline{n_e}/\overline{r^2}\ \overline{n_e r^2}$ will be approximately 10. Hence, for $r = 1$, an upper limit for the intensity of λ 5876 is 0·1 rayleigh and an upper limit for the intensity of λ 10 829 is 0·5 rayleighs.

From Fig. 5 we deduce that $\overline{n_e r^2}$ is less than 17 or 800 for temperatures of 500 or 2×10^5 °K, respectively, at $r = 1$ A.U.

10. CONCLUSION

If the neutral hydrogen density, n_H, electron density, n_e, and kinetic temperature, T_e of the

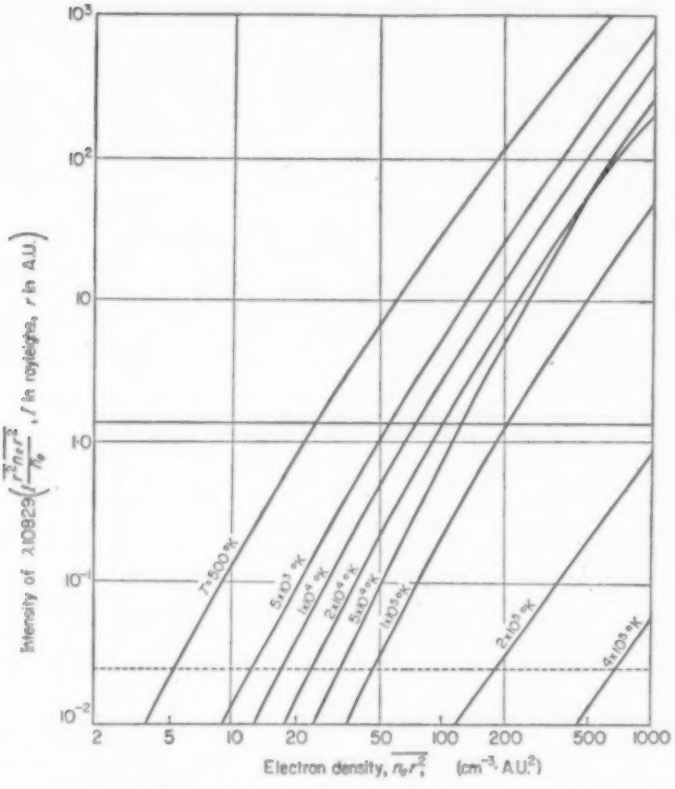


Fig. 5. The calculated intensities of the helium emission line at 10829A from a static interplanetary gas. The horizontal solid line is the observed upper limit of the emission intensity normalized to r=1 A.U. as explained in the text. The horizontal dashed line is the approximate limit of observationdue to the zodiacal light continuum at an elongation of 30°. This too is normalized to r=1 A.U.

interplanetary gas can be independently measured, then one can infer whether the interplanetary gas (if it actually exists) is static or non-static. n_e can be obtained from measurements made by Blackwell and Ingham (unpublished); n_{II} and T_e can be obtained from an emission line. It is possible that Hα may be present in the zodiacal light, and its observation may, with some difficulty, give a Doppler temperature. It is also possible that $\lambda\lambda$ 5876 and 10,820 Å of HE I are weakly present in the zodiacal light to give, again with some difficulty, measurements of n_{II} and T_e. The intensity of λ 10,829 is above the zodiacal light continuum intensity for static electron densities below 50 cm^{-3} and kinetic temperature below 10^5 °K. For higher temperatures, the electron density must be higher if λ 10,829 is to be observed. The numerical results are obtained by assuming total equilibrium in the gas, which must therefore exist for several centuries. If the gas is moving away from or towards the sun at a rate which competes with the rate at which equilibrium is obtained, there will be an excess

of ions or atoms and the intensities of the emisson lines will be, respectively, less or more than calculated.

As previously mentioned, the accuracy of the calculated intensities and ionizations depends on the accuracy of the assumed solar radiation near the ionization limits of H I, He I and He II at 900, 500 and 200 Å. The observation of the He I lines and of Hα is difficult because of their low intensities and the presence of intense OH bands in the night sky spectrum.

Acknowledgements—The author wishes to acknowledge many helpful discussions with Dr. D. E. Blackwell and to thank him and Mr. M. F. Ingham for allowing reference to their unpublished data. The work was supported by a Post-Doctorate Overseas Fellowship from the Canadian N.R.C.

REFERENCES

1. F. L. Whipple and J. L. Gossner, *Astrophys. J.* **109**, 380 (1949).
2. W. R. Piggott and C. W. Allen, *9e Rapport de la Commission pour l'étude des relations entre les phénomènes solares et terrestres* p. 77 (1957).
3. M. F. Smyth, *Mon. Not. R. Astr. Soc.* **114**, 137 (1954). *Mon. Not. R. Astr. Soc.* **114**, 503 (1954).
4. A. Vallance Jones, private communication (1959).
5. E. J. Opik, *Z. Astrophys.* **35**, 43 (1954).
6. H. Elsässer, *Mitt. Astr. Ges. Hamburg*, 61 (1958).
7. J. E. Kupperian, Jr., E. Bryam, T. A. Chubb and H. Friedman, *Planet. Space Sci.* **1**, 3 (1959).
8. J. T. Jeffries and R. G. Giovonelli, *Aust. J. Phys.* **7**, 574 (1954).
9. J. T. Jeffries, *Aust. J. Phys.* **6**, 22 (1953).
10. H. Zirin, *Astrophys. J.* **123**, 536 (1956).
11. J. B. Zirker, *Astrophys. J.* **129**, 424 (1959).
12. L. H. Aller, *Gaseous Nebulae*. Chapman and Hall, London (1956).
13. A. Dalgarno and M. R. C. McDowell, *The Airglow and the Aurorae*. A Symposium held at Belfast in September 1955. (Edited by A. Dalgarno and E. B. Armstrong), p. 340. Pergamon Press, London (1959).
14. V. A. Ambartsumyan, *Theoretical Astrophysics*. Pergamon Press, London (1958).
15. A. Unsöld, *Physik der Sternatmosphären*. Springer, Berlin (1955).
16. C. W. Allen, *Astrophysical Quantities*. Athlone Press, London (1955).
17. R. v. d. R. Woolley and C. W. Allen, *Mon. Not. R. Astr. Soc.* **108**, 292 (1948).
18. G. Cillié, *Mon. Not. R. Astr. Soc.* **92**, 820 (1932).
19. L. Goldberg, *Astrophys. J.* **90**, 414 (1939).
20. C. de Jager, *Ann. Géophys.* **11**, 330 (1955).
21. H. Friedman, *9e Rapport de la Commission pour l'étude des relations entre les phénomènes solares et terrestres*, p. 41 (1957).
22. J. J. Hopfield, *Astrophys. J.* **72**, 133 (1930).
23. J. E. Kupperian, Jr., E. Bryam, T. A. Chubb and H. Friedman, *I.G.Y. Rockets Rep. Ser. No. 1*, July 30 (1958).
24. O. Struve, *Sky and Telescope*, p. 445 (1958).
25. I. S. Shklovsky, *Paper presented at 5th General Assembly C.S.A.G.I.*, Moscow (1958).
26. D. M. Hunten, F. E. Roach and J. W. Chamberlain, *J. Atmos. Terr. Phys.* **8**, 345 (1956).
27. J. E. Kupperian, Jr., E. Byram, T. A. Chubb and H. Friedman, *Ann. Géophys.* **14**, 329 (1958).
28. C. T. Elvey, *Astrophys. J.* **111**, 432 (1950).
29. A. B. Meinel, *Astrophys. J.* **111**, 433 (1950).
30. I. S. Shklovsky, *Ann. Géophys.* **14**, 414 (1958).
31. D. E. Blackwell, *Mon. Not. R. Astr. Soc.* **115**, 629 (1955).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 99–103. Printed in Great Britain.

MAGNETIC STORMS AND COSMIC RADIO NOISE ON 25 Mc/s AT AHMEDABAD (23°02′N; 72°38′E)

R. V. BHONSLE and K. R. RAMANATHAN
Physical Research Laboratory, Ahmedabad 9, India

(*Received 5 September* 1959)

Abstract—Analysing the intensities of cosmic radio noise on 25 Mc/s at Ahmedabad, it was found that during days of magnetic storms there was, in addition to the usual variation of intensity due to the diurnal changes in *D*-region ionisation and *F*-region absorption, a further variation. This showed itself as a decrease in the attenuation on the first two days of the storm and an increase in the attenuation on the third day. The decreased attenuation is most clearly seen in the first half of the night on the first and second days after the commencement of a storm and is associated with a decrease in *F*-scatter.

In a recent paper, Kotadia and Ramanathan[1] have discussed the effects of geomagnetic disturbances on the behaviour of the ionosphere over Ahmedabad. In low latitudes, the *F2*-layer of the ionosphere gets much more disturbed than the lower ionospheric layers. The critical frequency of the *F2*-layer, the height of maximum electron density (h_mF2) and *F*-scatter all undergo significant changes. In auroral latitudes, intense absorption of radio waves below the *E*-region of the atmosphere is usually associated with magnetic storms.

An analysis of the data of cosmic radio noise on 25 Mc/s at Ahmedabad over a period of one year from March 1957 to February 1958 showed that the total ionospheric attenuation suffered by cosmic radio noise could be divided into two components, a *D*-region component and an *F2*-region component. The *D*-region component normally depends on the sun's zenith distance and shows maximum in summer and minimum in winter. The *F2*-region component, when f_0F2 exceeds a certain minimum value, increases rapidly with increasing f_0F2. Ramanathan and Bhonsle[2] have shown that at least in the early part of the night, *F*-scatter is also an important factor which affects night-time attenuation. Since f_0F2 and *F*-scatter change appreciably on magnetically disturbed days, it was felt that an analysis of the data of cosmic radio noise absorption under disturbed ionospheric conditions would give useful information regarding the ionosphere.

The present paper deals with the effect of a few selected magnetic storms on the total attenuation of cosmic radio noise on 25 Mc/s at Ahmedabad. The list of magnetic storms used for this study are given in Table 1.

Table 1
List of Eleven Magnetic Storms all with Sudden Commencement (1957-58)*

Date of beginning	Time of beg.n-ning	Date of end	Time of end	Intensity
1957				
3 Aug.	20.58	4	14	m
6 Aug.	10.08	6	20	m
30 Aug.	00.20	31	01	m
2 Sept.	08.15	4	00	s
4 Sept.	18.00	6	04	s
13 Sept.	05.49	13	21	s
21 Sept.	15.06	22	16	s
22 Sept.	18.45	24	02	s
29 Sept.	05.15	30	16	s
26 Nov.	06.52	27	23	ms
1958				
11 Feb.	06.24	12	18	s

*Time in 75° EMT

When making a day to day study of cosmic radio noise, it was found that on some days, the total attenuation had abnormally low values throughout the day and till about midnight. This often happened for two days in succession. On the third day, the attenuation showed a tendency to rise above the normal, after which it returned to normal. It was also found that this feature was associated with magnetic storms. Fig. 1 shows the hourly attenuations of cosmic radio noise on 25 Mc/s at Ahmedabad during six magnetically disturbed periods in 1957–58 superposed on the month's quiet day attenuations for each period. The magnetic storm which occurred on 13 September 1957 was of comparatively short duration and an apparent exception. In general, a decrease in the attenuation of cosmic radio noise occurred on the first two days of the storm and excess attenuation on the third day.

Fig. 2 shows hourly attenuations of cosmic

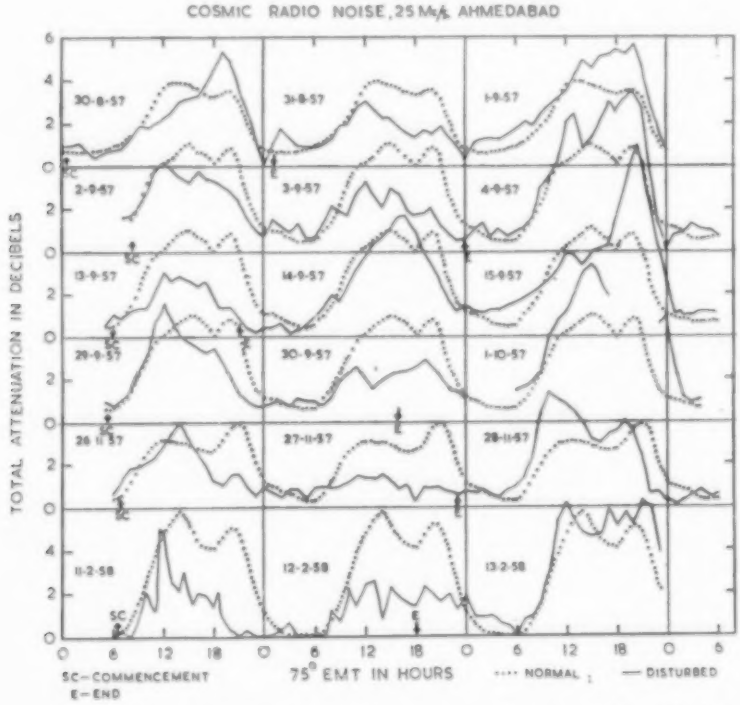


Fig. 1. Hourly attenuation of C.R. noise on 25 Mc/s at Ahmedabad in six magnetically disturbed periods in 1957-58 together with the month's quiet day attenuation for each period. The times of commencement of and end of each magnetic storm are marked in the figure.

radio noise in three periods in 1957 when the magnetic disturbance was more complex and more than one commencement could be identified. Under such conditions, the variation of total attenuation of cosmic radio noise on 25 Mc/s does not follow a fixed pattern.

The deviations of the total attenuation of cosmic radio noise from the corresponding monthly mean gives the effect of a magnetic disturbance on the ionosphere. The usual decrease or increase in the total attenuation of cosmic radio noise associated with changes in

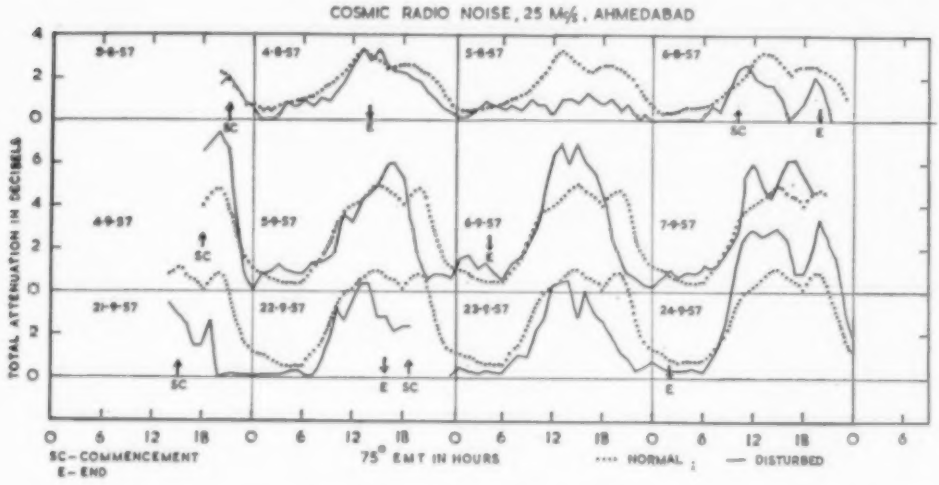


Fig. 2. Hourly attenuation of C.R. noise in three periods in 1957-58 when the magnetic disturbance was complex and more than one commencement could be identified.

f_0F2 when it exceeded a certain value was always noticed. But on magnetically disturbed days there was an additional decrease in attenuation which depended on local time. To determine the average diurnal variation of the change in attenuation due to magnetic storms, the 11 magnetic storms given in Table 1 were analysed. Fig. 3 gives the average hourly values of the departure from normal of cosmic radio noise on the first two days after the commencement of a magnetic storm. It is clear that the effect of a magnetic storm is in general to reduce the attenuation throughout the day except for the period midnight to sunrise. During the later part of the night, low values of f_0F2 are usually observed and the change in f_0F2 has no effect. After sunrise, the departure of the total attenuation from normal goes on increasing till 16 hr local time. After a temporary decrease which lasts till 18 hr, it again increases till 20 hr when it reaches a maximum. It returns to normal by about 2 hr. From 2 to 5 hr the effect of the storm is negligible. The disturbance diurnal variation of the total attenuation of cosmic radio noise described above can be correlated with the results obtained earlier by Kotadia and Ramanathan.[1] They showed that the largest changes in the ionosphere occur on the same day, or the day following the commencement of the magnetic storm. The change at Ahmedabad is a daytime decrease in f_0F2 in contrast to the changes at equatorial stations. The occurrence of spread-F at Ahmedabad is mainly a night-time feature. It was found that during magnetic disturbances, spread-F was generally absent or small, and the usual increase in height and in f_0F2 at 20-21 hr did not occur. It begins usually after 19 hr. In

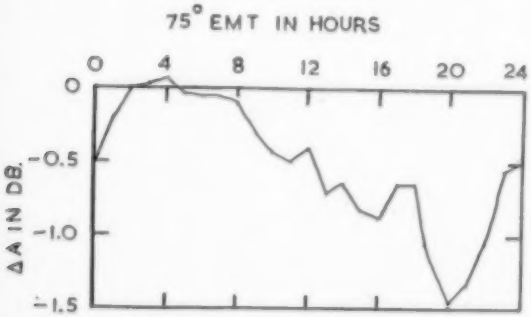


Fig. 3. Average diurnal variation of the departure from normal of C.R. noise on the first two days of a magnetic storm.

view of the above facts, during daytime, only the changes in f_0F2 are likely to be important in influencing the attenuation of cosmic radio noise. After sunset, however, both the change in f_0F2 and the absence of F-scatter on magnetically disturbed nights would be jointly responsible for the observed rapid decrease in total attenuation.

From the study of the disturbance diurnal variation of the individual storms, it was seen that, in many cases, after the magnetic storm ended, there occurred on the following day an increase in the total attenuation above the normal. Fig. 4 shows the average hourly attenuations of cosmic radio noise associated with magnetic disturbances (mean of five periods). Zero hour has been taken as 0 hr LMT of the day on which the cosmic radio noise attenuation increased above normal at the end of a magnetic storm. For two days before the zero hour, an overall decrease in the total attenuation is observed following the general pattern shown in Fig. 3. On the first day after the zero hour an overall increase in the total attenuation is observed. It is interesting that on this day, the maximum increase in the total attenuation occurs at 20–21 hr. On the two previous nights, at these same times, minimum total attenuation is recorded. This observed overall increase in the total attenuation after the end of a magnetic storm may be partly due to a rise in f_0F2 above normal and partly to a tendency for increased night-time F-scatter. Nothing particular happens on the next two days.

CONCLUSION

The geomagnetic effects on the ionosphere over Ahmedabad are discussed with the aid of cosmic radio noise data on 25 Mc/s. Since $F2$-attenuation depends on the nearness of f_0F2 to the exploring frequency i.e. 25 Mc/s, this method provides a sensitive method of studying the effects of magnetic storms on the $F2$-layer as a whole. The disturbance diurnal variation of the total attenuation of cosmic radio noise shows a dependence on local time. The maximum decrease in total attenuation occurs at about 20 hr local time. An explanation of the observed variation is given in terms of the known variations of f_0F2 and F-scatter on magnetically disturbed days. After the end of the storm, an overall increase in attenuation is observed, which lasts for a day.

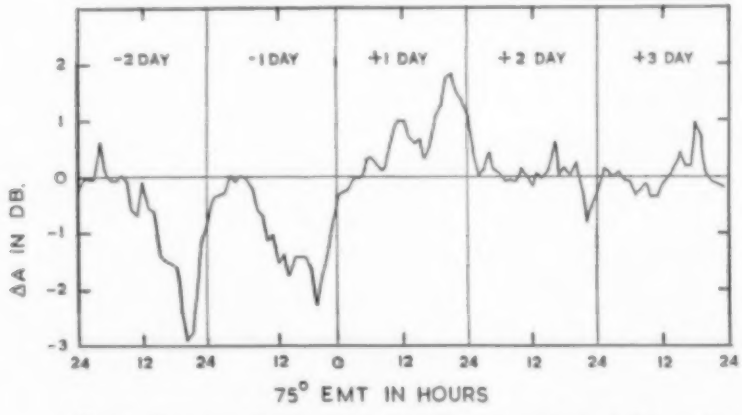


Fig. 4. Departure from normal of hourly attenuations of C.R. noise associated with magnetic disturbances (mean of five periods). Zero hour has been taken as 0 hr LMT of the day on which the C.R. attenuation increased above normal at the end of the magnetic storm. This generally happened two days after the commencement of a simple storm.

Acknowledgements—The thanks of the authors are due to the Indian Council of Scientific and Industrial Research for supporting the work. One of the authors (R.V.B.) was in receipt of a Senior Fellowship from the Council during part of the period of the work.

REFERENCES

1. K. M. KOTADIA and K. R. RAMANATHAN, *Magnetic and ionospheric disturbances in low latitudes*, to be published.
2. K. R. RAMANATHAN and R. V. BHONSLE, Cosmic radio noise absorption on 25 Mc/s and *F*-scatter, *J. Geophys. Res.* **64**, 1635 (1959).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 104–109. Printed in Great Britain.

SOLAR RADIO BURSTS AND COSMIC RAYS

A. R. THOMPSON and A. MAXWELL

Radio Astronomy Station of Harvard College Observatory, Fort Davis, Texas

(*Received* 24 *September* 1959)

Abstract—Current theories of the origin of solar radio bursts suggest that fast drift (Type III) bursts may be generated by the emission of solar particles with velocities approaching those of cosmic rays. An examination of cosmic ray intensity and solar radio data over the two-year period 1956 October–1958 September shows, however, no correlation between fast drift bursts and cosmic ray increases on the earth. Solar flares, with which fast drift bursts are associated, also appear to be unrelated to cosmic ray increases. On the other hand there have been a number of cases of increases in low energy cosmic ray particles, recorded by equipment in satellites and balloons, which were probably related to outbursts of solar continuum (Type IV) radiation. Forbush type decreases in cosmic ray intensity, and accompanying magnetic storms, are regularly preceded by slow drift (Type II) radio bursts.

1. INTRODUCTION

This paper is concerned with the association between solar radio bursts and variations in cosmic ray intensity. The radio data used in the investigations were recorded at the Harvard Radio Astronomy Station, Fort Davis, Texas. Most of the data were obtained during the period 1956 October through 1958 September, and were recorded on sweep frequency equipment covering the range 100–580 Mc/s.[1] Some more recent observations covering the range 25–580 Mc/s are also used. The cosmic ray data were obtained from the hourly counts of the neutron monitor pile at Climax, Colorado. Cosmic ray counters measuring local neutron production respond to primary particles in the rigidity range from 10^9 V upwards, which contains over 95 per cent of all primary cosmic radiation.[2] Zurich provisional sunspot numbers and flare data were obtained from the *C.R.P.L. Mon. Bull. Solar-Geophys. Data* (Boulder, Colorado).

Three spectral types of radio bursts are considered separately. In Section 2 we examine the possibility that fast drift (Type III) radio bursts originate during the emission of cosmic ray particles from the sun, and compare observed values of the intensities of cosmic rays and fast drift bursts over a two year period at solar maximum. The association of cosmic rays with sunspot numbers is also examined in this section. Section 3 concerns the association of cosmic ray intensity with solar flares; this is of importance since several types of radio bursts are associated with flares. Section 4 comprises a brief discussion of the relation between Forbush-type decreases in cosmic ray intensity and the occurrence of slow drift (Type II) radio bursts that are also associated with geomagnetic storms. Section 5 is concerned with a number of cases in which increases in low energy particles occurring at high altitudes have been accompanied by outbursts of solar continuum radiation (Type IV). The physical implications of some of the results are discussed in the final section.

2. FAST DRIFT RADIO BURSTS

We first examine the suggestion that fast drift radio bursts are generated by the emission of cosmic rays from the sun. Two groups of fast drift bursts recorded at Fort Davis are shown in Fig. 1(a), from which it will be seen that each burst has a lifetime of several seconds, during which the frequency of emission drifts rapidly towards the lower end of the radio spectrum.

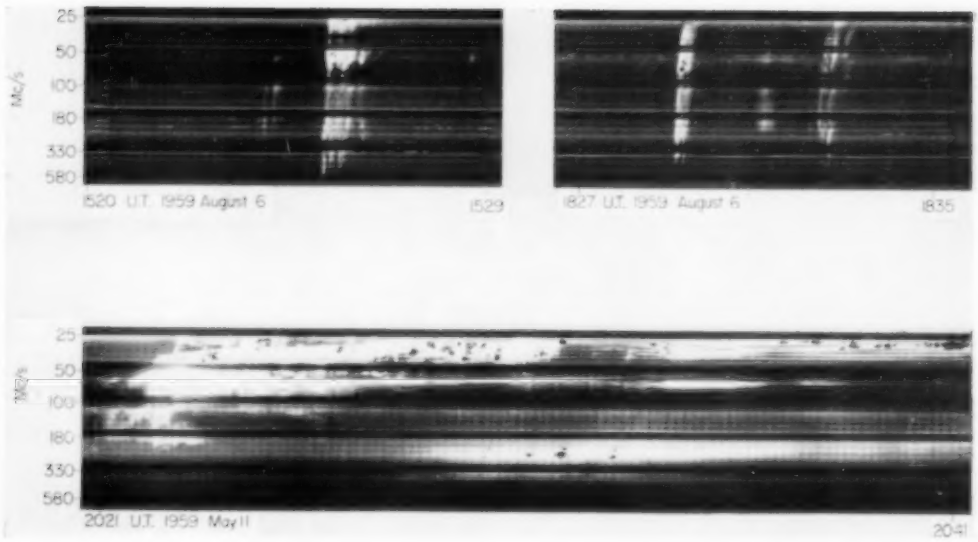


Fig. 1. (a) Upper records: fast drift bursts (Type III). (b) Lower record: slow drift burst (Type II), followed by continuum radiation (Type IV) over the range 180-400 Mc/s.

[facing page 104]

VOL
2
1959

The bursts are commonly attributed to a disturbance that moves outwards through the solar atmosphere, and excites emission from levels of progressively lower plasma frequency. From the rate of change of frequency, it is estimated that the velocity of the disturbing agency moving outwards through the solar corona is about 100 000 km/sec.[3] Wild, Roberts and Murray[4] have suggested that this primary disturbance might be cosmic ray particles being emitted from the sun. It must immediately be remarked that a velocity of 100 000 km/sec is considerably lower than the velocity usually attributed to cosmic ray particles, and that it corresponds to a time of travel of approximately 25 min for a direct path between the sun and the earth. In this connection, Wild and his colleagues pointed out that, in the few instances when an intense flare has been accompanied by a great increase in cosmic rays, the observed time delay between the flare and the cosmic ray increase was of this order.

To examine the validity of the above suggestion, we have compared daily average figures of fast drift solar radio bursts and cosmic ray intensity over the two years 1956 October 1 through 1958 September 30. During this period more than 25 000 fast drift bursts were recorded at Fort Davis, a total that corresponds to a mean rate of 3·3 fast drift bursts/hr. Examination of the data also shows that the emission of fast drift bursts is often enhanced for periods of 2 or 3 days, during which the hourly rate may be as much as 10 times the mean figure for the 2-year period. A plot of the daily figures for fast drift bursts with daily mean values of cosmic ray intensity given by Thompson[5] reveals no evidence of a positive association between the two phenomena, although some periods of high radio emission coincide with the 27 day recurrent decreases in cosmic ray intensity. This effect is by no means well defined over the 2-year period, and in order to settle the question unambiguously, cross correlation coefficients between the fast drift bursts and cosmic rays have been calculated with an *IBM* 650 computer. The daily figures of cosmic ray intensity were averaged over a complete day, 0–24 hr U.T.; the fast drift bursts were weighted according to intensity and averaged over the observing period, approximately 12–24 hr U.T., of each day. As a control index, cross correlation coefficients between cosmic ray intensity and sunspot number were also computed for the 2-year period. The coefficients were taken for a range of time displacements of up to ±14 days, and the results are plotted in Fig. 2.

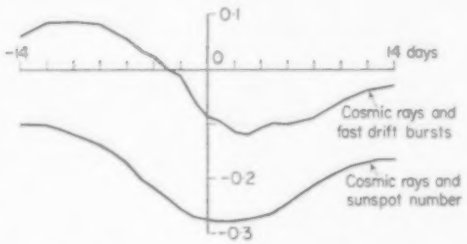


Fig. 2. Cross correlation coefficients.

Positive values of displacement indicate a comparison of cosmic ray data with solar data recorded a given number of days earlier. The figure shows that the cross correlation curve for cosmic rays and fast drift bursts has a tendency towards negative values. The curve for cosmic rays and sunspot number is entirely negative.

The significance of cross correlation coefficients depends upon the effective number of independent values from which they are derived. As shown by the autocorrelation coefficients plotted in Fig. 3, the day to day values of cosmic ray intensity are not independent; and the same situation obtains for the fast drift bursts and the sunspot numbers. The autocorrelation coefficients in fact indicate that the effective number of independent values for the cosmic ray data for the 2-year period is about 100, for the fast drift bursts about 200, and for the sunspot numbers about 100. For 100 samples, the value of the cross correlation coefficient corresponding to 0·05 probability of chance association is 0·19.[6] Thus the cross correlation values for cosmic rays and fast drift

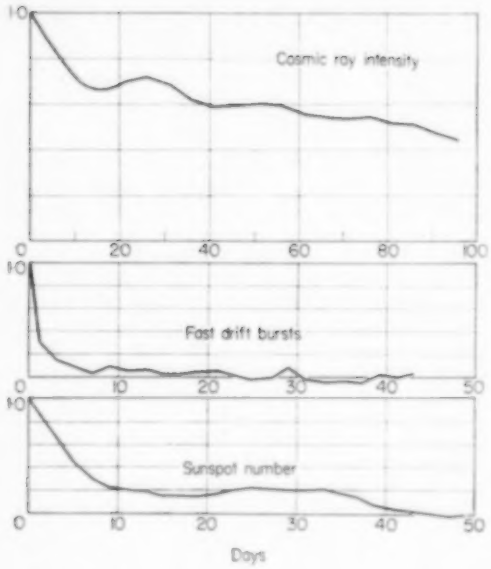


Fig. 3. Autocorrelation coefficients: note the increases at 27 days, appropriate to the synodic rotation period of the sun.

bursts are not significant. This result lends no support to the suggestion that fast drift bursts originate from the emission of cosmic rays from the sun.

Eighty of the larger groups of fast drift bursts, each of high intensity, containing more than 10 bursts per group, and isolated by at least ± 1 hr from other groups of fast drift bursts, were also compared with cosmic ray figures. This was done with a superposed epoch diagram, not reproduced here, for the neutron counts over a period ± 12 hr centered about the time of the individual groups of fast drift bursts. It was not possible to draw any conclusions from the analysis because the large residual effect of the diurnal variation in the cosmic ray figures masked all effects less than 2 per cent in amplitude.

In the case of cosmic rays and sunspot numbers the curve of correlation coefficients reaches a value of $-0{\cdot}28$, which is statistically significant. This suggests that the effect producing the long term decrease in cosmic ray intensity over the maximum of the solar cycle[7] is also responsible for decreases over restricted periods of a few days. This result is in accord with the observation of Venkatesen[8] that an association exists between transient decreases in cosmic ray intensity and the central meridian passage of active solar regions. Present theories attribute the negative modulation of cosmic ray intensity to the diffusion of extra-solar cosmic rays through magnetic clouds existing in the interplanetary medium.

Correlation coefficients were also computed for cosmic rays and daily averages of noise storm intensity (radio spectral Type I) for the period 1956 October 1 through 1958 September 30. The resulting curve was very similar to that for cosmic rays and fast drift bursts, with no values above the level of significance.

3. SOLAR FLARES

A positive association between cosmic ray intensity and the occurrence of sixteen solar flares for the years 1951 through 1953 has been reported by Firor.[9] Since fast drift radio bursts are also associated with flares,[1, 10] this would suggest that some small positive association may exist between the fast drift bursts and cosmic ray intensity which has not been revealed by the above analysis. To check this point, an investigation has been made of the association of flares with cosmic ray intensity for the year 1956 October through 1957 September. This study was carried out in a manner analogous to that employed by Firor. Flares of importance $\geqslant 2$-, falling within the morning time zone for Climax, Colorado, were used. The result is plotted in Fig. 4, in which the abscissa shows time relative to the flares, and the ordinate gives cosmic ray data from the Climax monitor pile averaged over 76 flares. The figure shows that, for the 12 month period under consideration, there was no association between the cosmic ray count and solar flares. The shape of the curve in Fig. 4 is attributed wholly to the residual diurnal variation in the neutron count. This result, derived from data taken at sunspot maximum, contrasts with that of Firor[9]: for 16 flares at sunspot minimum Firor found a

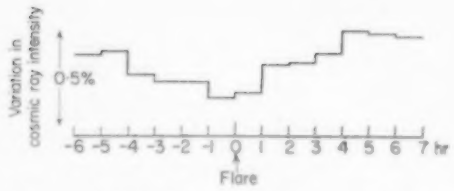


Fig. 4. Superposed epoch diagram of cosmic ray intensity for seventy-six flares.

sharp increase in cosmic rays, of amplitude about 1 per cent and duration 2 hr.

4. SLOW DRIFT RADIO BURSTS

In an earlier paper[11] it has been shown that there is a significant association between the occurrence of the slow drift type of solar radio bursts and of terrestrial magnetic storms. For the two year period 1956 October 1 through 1958 September 30, at least 60 per cent of magnetic storms were preceded by slow drift bursts. The mean time delay between the radio bursts and magnetic storms was 33 hr.

Forbush[12] has shown that decreases in cosmic ray intensity accompany many intense magnetic storms. It would therefore be inferred that a statistical association should exist between slow drift radio bursts and decreases in cosmic ray intensity of the type described by Forbush. The present investigation confirms this association. There were 26 well defined cosmic ray decreases of the Forbush type recorded at the neutron monitor pile at Climax, Colorado, during the period 1956 October 1 through 1958 September 30. At least 56 per cent of these decreases were preceded by a slow drift burst within 4 days (the association would probably be higher if complete 24 hr radio spectral data were available), and the mean time delay between the radio bursts and the Forbush decreases was 35 hr.

5. CONTINUUM RADIATION AND SOLAR COSMIC RAYS

Recent observations, made with scintillation and Geiger counters carried in satellites[13] and balloons,[14, 15] have shown that protons with energies up to a few hundred MeV quite frequently bombard the upper atmosphere at high magnetic latitudes for some days following large solar flares. It has been suggested[13] that these solar protons are low-energy examples of the great cosmic ray outbursts from the sun, such as occurred on 1956 February 23. It is of considerable interest to compare these observations of low-energy cosmic rays in the outer atmosphere with the radio spectral observations. The relevant data are summarised in Table 1. It will be seen at once that for all cases where radio data were available, the solar protons were preceded or accompanied by large solar radio bursts of continuum radiation of the type normally referred to as spectral Type IV. An example of this type of burst is given in Fig. 1(b). A continuum burst is characterised by an enhancement of solar radiation lasting from 5 min to over 12 hr. The radiation is normally very steady and is often intense. In the frequency range 25 to 580 Mc/s it may cover a band several hundred megacycles wide, and it sometimes drifts towards lower frequencies over a period of the order of an hour. Such outbursts are recorded at Fort Davis on about 15 days per year at solar maximum. Boischot and Denisse[16] have suggested that this type of radio emission arises from synchrotron radiation in the solar corona.

Each of the cases cited in the table also coincided with an increase in ionospheric absorption of a special type* described by Reid and Leinbach[17]. These absorption events are confined to polar latitudes and appear as an increase in ionospheric absorption of cosmic radio waves. They are attributed to *D*-region ionization caused by solar protons with velocities a few tenths of the velocity of light. Such absorption events occur about 10 times per year at solar maximum, and usually are preceded by a major flare and an outburst of continuum radiation.

As shown in Table 1, the complete sequence of events comprises a flare of high importance, followed approximately 15 min later by an

* Ionospheric polar cap absorption, or "Type III" absorption.

Table 1. Continuum Radiation (Type IV) and Solar Cosmic Rays 1958 Jan. 1 to 1959 July 31

Date	Flare		Continuum radiation (Type IV)		Solar protons		
					Indirect observations (polar ionospheric absorption)		Direct observations (satellites, balloons)
	Start (U.T.)	Importance	Start (U.T.)	Duration (hr)	Start (U.T.)	Duration (hr)	
1958 Feb. 9	21.08	2+	21.16	1·8	<07.00 Feb. 10	>30	
Mar. 25			outside observing hr		<22.30	>96	Balloon[13]
Apr. 10	<10·10	2	outside observing hr		11·30	>40	
July 7	00.22	3	00.52	>0·9	01.30	>120	
July 29	<03.03	3	03.21*	>1·7	04.05	>30	
Aug. 16	<04.32	3	outside observing hr		06.00	>60	Satellite[13]
Aug. 22	14.17	3	15.00†	>10·3	<17.00	>80	Satellite[13], Balloon[14]
Aug. 26	00.05	3	00.20	>0·9	01.00	>89	Satellite[13]
Sept. 22			outside observing hr		14.30	>82	Balloon[20]
1959 May 10	21.02	3+	21.16	>4·3	01.30 May 11	>200	Balloon[21]
	<23.10	3					
May 11	20.06	2+	20.28	0·3			
July 10	<05.14	3	outside observing hr		07.00	>90	Balloon[20]
July 13	19.20	2	19.37	0·5	<07.00 July 14	>51	Balloon[20]
July 14	<03.25	3	04.01*	>2·2			Balloon[20]
July 16	21.14	3	21.21	>4·3	<22.50	>34	

* Data from 40–240 Mc/s equipment at Sydney, Australia.

† Approximate starting time: first 30 min inferred from 60 Mc/s fixed frequency equipment.

outburst of continuum radiation; about 1 hr later an enhanced flux of solar protons reaches the earth.[13]

Further investigation shows that solar protons in the high atmosphere, observed directly with satellite or balloon equipment or indirectly by polar radio absorption, are not accompanied by a significant increase in the hourly rate of fast drift radio bursts. Also, the great majority of flares associated with the emission of solar protons are not accompanied by slow drift radio bursts.

6. CONCLUSION

In reviewing the above evidence we conclude that, if the fast drift radio bursts are generated by cosmic rays leaving the sun, then these cosmic rays are not detected by existing neutron counters on the earth. It is possible that such cosmic rays escape detection because of their small number. But the existing data concerning the fast drift bursts strongly suggest that the primary disturbance that causes the radio bursts has a velocity of only 10^5 km/sec, that is, 0·3 times the velocity of light. The radio evidence also shows that the primary disturbance is sometimes greatly decelerated or even brought to rest. Thus the particles, if they manage to escape from the solar atmosphere, will be of low rigidity and would not penetrate the terrestrial atmosphere except, possibly, at the magnetic poles; they should, of course, be detectable with sensitive equipment carried in a satellite. Alternatively, it may be that the fast drift bursts are caused not by fast travelling corpuscular matter, but by shock waves, as suggested by Krook[19].

In a limited number of cases, great outbursts

of continuum radiation have preceded or coincided with the bombardment of the upper atmosphere by solar protons. The continuum radiation is likewise associated with the occurrence of a certain type of radio absorption in the polar ionosphere which is attributed to solar particles of similar energy. Thus, there seems little doubt that a physical connection exists between the emission of protons and of continuum radiation from the sun. Both phenomena may last for many hours after the end of the associated solar flare. The energy of the protons estimated from balloon and satellite observations is in the range 30–300 MeV, which corresponds to velocities of the order of 0·2–0·7 times the velocity of light. The electrons that produce the continuum radiation by synchrotron emission probably have velocities of the same order.

Acknowledgements—The work described in this paper is part of a program supported by the Air Force Cambridge Research Center through the Sacramento Peak Observatory. The authors gratefully acknowledge the assistance of Mr. R. S. Lawrence of the National Bureau of Standards, Boulder, who programmed the correlation computations. They also wish to thank Prof. J. A. Simpson of the University of Chicago for data from the neutron monitor pile at Climax, Colorado, Mr. H. Leinbach of the Geophysical Institute, Alaska, for unpublished data on ionospheric absorption, and Mr. G. Garmire for his assistance at the Station.

REFERENCES

1. A. Maxwell, G. Swarup and A. R. Thompson, *Proc. Inst. Radio Engrs.* **46**, 142 (1958).
2. J. A. Simpson, *Proc. Nat. Acad. Sci., Wash.* **43**, 42 (1957).
3. A. Maxwell and A. R. Thompson, in preparation.
4. J. P. Wild, J. A. Roberts and J. D. Murray, *Nature, Lond.* **173**, 532 (1954).
5. A. R. Thompson, *Paris Symposium on Radio Astronomy*, p. 210. Stanford University Press (1959).
6. R. A. Fisher, *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh (1954).
7. S. E. Forbush, *J. Geophys. Res.* **63**, 651 (1958).
8. D. Venkatesan, *J. Geophys. Res.* **64**, 505 (1959).
9. J. Firor, *Phys. Rev.* **94**, 1017 (1954).
10. G. Sharup, P. H. Stone and A. Maxwell, *Astrophys. J.*, in the press.
11. A. Maxwell, A. R. Thompson, and G. Garmire, *Planet. Space Sci.* **1**, 325 (1959).
12. S. E. Forbush, *Terr. Mag.* **43**, 203 (1938).
13. P. Rothwell and C. McIlwain, *Nature, Lond.* **184**, 138 (1959).
14. K. A. Anderson, *Phys. Rev. Letters* **1**, 335 (1958).
15. P. S. Freier, E. P. Ney and J. R. Winckler, *J. Geophys. Res.* **64**, 685 (1959).
16. A. Boischot and J. F. Denisse, *C.R. Acad. Sci., Paris* **245**, 2194 (1957).
17. G. C. Reid and H. Leinbach, *J. Geophys, Res* **64**, 1801 (1959).
18. A. R. Thompson and A. Maxwell, *Nature, Lond.* in press.
19. M. Krook, private communication.
20. J. R. Winckler, private communication.
21. E. P. Ney, J. R. Winckler and P. S. Freier, *Phys. Rev. Letters* **3**, 183 (1959).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 110–120. Printed in Great Britain.

RADIO WAVE PROPAGATION CHARACTERISTICS OF A SIMPLE IONOSPHERIC MODEL BASED ON ROCKET DATA

L. ZHEKULIN

Academy of Sciences of the U.S.S.R., Moscow

(*Received* 17 *January* 1959)

Abstract—In the U.S.S.R. and abroad in the last few years valuable experimental data have been obtained with the help of rockets, relating to investigations of distribution of ionization with altitude in high atmospheric regions. The available results make it possible to attain a more specific understanding of ionosphere structure and lead to a number of conclusions pertaining to the problem of radiowave propagation at high altitudes.

1. IONIZATION DENSITY DISTRIBUTION WITH ALTITUDE

Results of measurements of effective electron concentration distribution in the ionosphere with the help of high-altitude geophysical rockets of the U.S.S.R. Academy of Sciences over the period of 1954–58 are given and analysed in the work of Gringauz[1]. The curve in Fig. 1 shows the dependence of effective electron concentration on an altitude on the basis of experiments carried out at 17.24 on 26 June 1954. Three diagrams in Fig. 2 give the results of experiments obtained in 1957. Rockets were launched at 6.18 on 16 May (diagram 1), at 6.27 on 25 August (diagram 2) and at 19.54 on 9 September (diagram 3). Now let us compare the given data with measurements made by Seddon *et al.*[2] Fig.

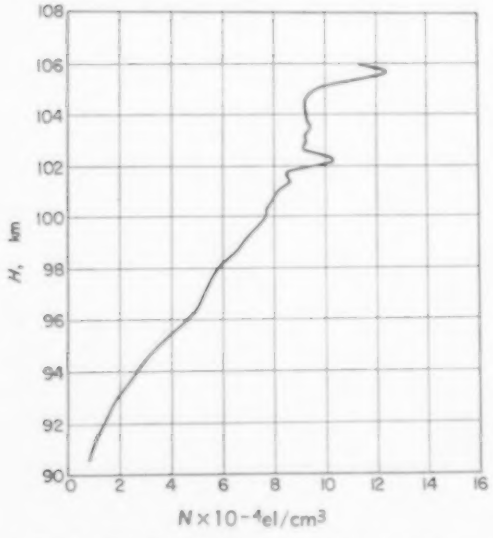


Fig. 1. The dependence of the effective electron concentration on an altitude (U.S.S.R. 1954).

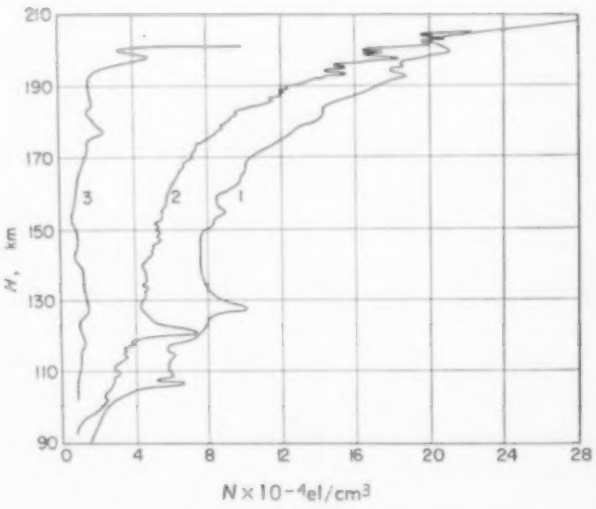


Fig. 2. The dependence of the effective electron concentration on an altitude (U.S.S.R. 1957).

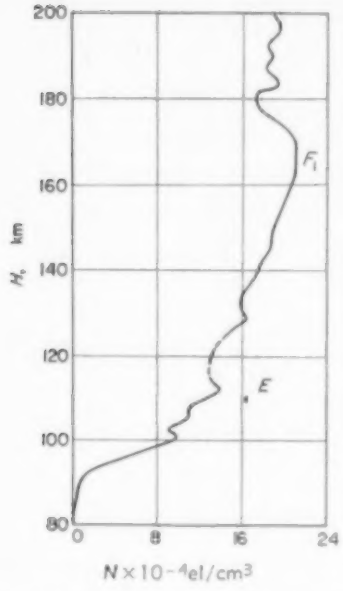


Fig. 3. The dependence of the effective electron concentration on an altitude (U.S.A. 1954).

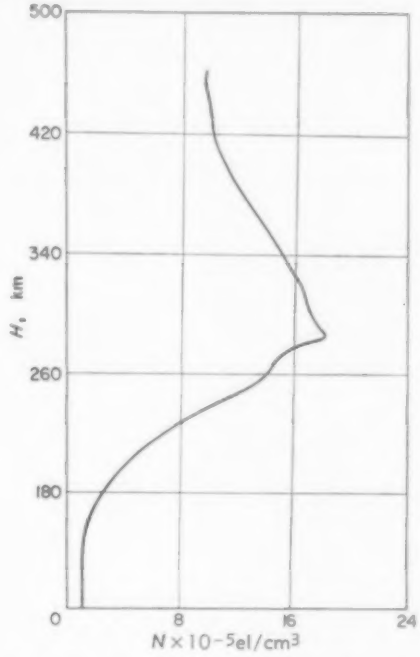


Fig. 4. The dependence of the effective electron concentration on an altitude (U.S.S.R. 1958).

3 represents ionization distribution with altitude, obtained at 10.00 on 7 May 1954, as a result of experiments with rocket Viking No. 10. Jackson[3] compared the curve shown in Fig. 3 with the distribution of electron density calculated with the use of altitude–frequency characteristics recorded during this rocket flight. The results turned out to be well in accord.

Consideration of the given diagrams in Fig. 1–3 shows that the results of the latest experiments obtained here and abroad are reconciled and enable us to conclude that the ionosphere at altitudes of 100–200 km appears to be a highly ionized medium, sharply outlined layers being perhaps absent; i.e. ionization changes from one ionosphere region to the adjacent without sharp alterations.

Now let us consider the ionization at high altitudes. Of special interest are observations made on 21 February 1958, during the flight of a Soviet geophysical rocket which reached an altitude of 473 km. Ionization distribution in the *F*-region of the ionosphere obtained as a result of observations carried out by means of this rocket is shown in Fig. 4. For comparisons, Fig. 5 illustrates the results of experiments[4] with the two-stage rocket Bumper Buck No. 5 (launched at 15.14 on 24 February 1949). Above the maximum, the experiments of foreign scientists showed a rapid decrease in ionization, and already at the altitude of $\approx$ 380 km the value turned out to be very low; while according to the U.S.S.R. data, the character of ionization decrease above the maximum was different: ionization slowly decreased with altitude retaining a high value of $\approx 10^6$ el/cm^3, even at an altitude of 470 km.

It is necessary to emphasize the value of investigations by means of rockets, which is that they provide a knowledge of the distribution of electron density as a function of true altitudes above the surface of the earth; while the data of ionospheric stations are referred to the effective altitudes, the difference between the latter and the true ones may amount in the

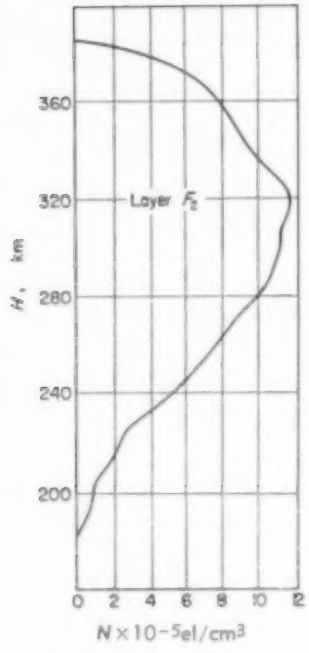


Fig. 5. The dependence of the effective electron concentration on an altitude (U.S.A., 1949).

F-region (depending on the state of the ionosphere) to 50–150 km.

Correct interpretation of ionosphere probing results is of important practical significance. The results of experiments with the Soviet artificial earth satellites pertaining to altitude-distribution of ionization are in qualitative agreement with the rocket data.

2. EFFECT OF IONIZATION DISTRIBUTION OBTAINED IN THE LATEST EXPERIMENTS ON RADIOWAVE PROPAGATION

Investigations of radiowave propagation in a heterogeneous ionized medium with ionization density distributions shown in the illustrated figures, will be based on Maxwell–Lorentz classical equations. Let us assume that a plane sinusoidal $\exp(j\omega t)$ electromagnetic wave falls normally on a plane interface of a heterogeneous ionized layer with a complex dielectric permeability (depending only on altitude Z):

$$\varepsilon^*(Z) = 1 - \frac{4\pi N(Z) \cdot e^2}{m\omega^2\left(1 - j\frac{V}{\omega}\right)} \qquad (1)$$

Neglecting the influence of the earth's magnetic field, we can write the basic differential equation of the problem as follows:

$$\frac{d^2E}{dZ^2} + \frac{\omega^2}{c^2} \cdot \varepsilon^*(Z)\, E = 0 \qquad (2)$$

At the beginning we shall consider the medium to be non-absorbing, and therefore in expression (1) the effective number of collisions, V, should be assumed to be zero. To provide the possibility of analytical examination of the formulated problem, the curves of effective electron density distribution, $N(Z)$, will be approximated by straight line sections. If we limit ourselves only to the main features of altitude-distribution of ionization it is enough to suggest that in region E

$$N = N_0 + K_1 Z \quad \text{for} \quad Z_1 \leqslant Z \leqslant 0$$

$$N = N_0 + K_2 Z \quad \text{for} \quad Z \geqslant 0$$

For instance, approximating the curve in Fig. 3 we shall have

$K_1 = 0{\cdot}113$; $K_2 = 0{\cdot}018$ and $N_0 = 10^5$ el/cm^3

For region F: $N = N_0 + K_1 Z \qquad Z_1 \leqslant Z \leqslant 0$

$$N = N_0 - K_2 Z \qquad Z_2 \geqslant Z \geqslant 0$$

The curve in Fig. 4 near the maximum is well enough approximated by a "triangular" layer with gradients

$K_1 = 0{\cdot}165$; $K_2 = 0{\cdot}0535$ and $N_0 = 1{\cdot}85 \times 10^6$ el/cm^3

By means of suitable substitution, equation (2) for a linear ionization law is converted to Airy's equation:

$$\frac{d^2E}{dt^2} = t \cdot E \qquad (2')$$

where in the absence of absorption a new variable is

$$t = -b_i^{1/3} \cdot \frac{\omega^2}{c^2}\varepsilon_i(Z);\ b_i = \left(\frac{mc^2}{4\pi c^2 K_i}\right)^2$$

$$\varepsilon_i = 1 - \frac{4\pi e^2}{m\omega^2}(N_0 \pm K_i Z) \qquad (3)$$

where $i=1, 2$. For a triangular layer in formula (3) the lower sign for $Z>0$ is taken, and $N_0=N_{max}$.

For Airy's functions $u(t)$ and $v(t)$ satisfying equation (2') we make use of determinations and designations introduced by Fok[5] (see Appendix).

In the problem considered, the advantage of using Airy's functions instead of Bessel's functions is that the former, and their first derivatives, preserve continuity at $t=0$.

Further, it is suggested that ionization density changes slowly over the distance of a wave length, and the continuity of tangential components of the electromagnetic field at planes $Z=Z_1$; $Z=0$, and $Z=Z_2$ for a triangular layer, and finiteness of the field intensity in the case of $Z\longrightarrow\infty$ are taken into consideration. Let us assume that the falling wave amplitude at the plane $Z=Z_1$ is equal to unity, then after the calculations we shall obtain (omitting time factor $\exp(j\omega t)$) the following expression for the strength of the resulting electric field at the interface $(Z=Z_1)$ of the ionized medium for the case of monotonic rise of ionization with altitude $(N=N_0+K_1Z)$.

$$E=1+\exp(j\,.\,2\alpha)\,.\,\exp\left(-j\frac{m\omega^2}{3\pi ce^2K_1}\right).\exp\left(j\frac{\pi}{2}\right) \tag{4}$$

for $Z=Z_1$ where

$$\operatorname{tg}\alpha=\frac{v'(t_0')\,v(t_0'')-\left(\frac{K_2}{K_1}\right)^{1/3}.\,v(t_0')\,v'(t_0'')}{u'(t_0')\,v(t_0'')-\left(\frac{K_2}{K_1}\right)^{1/3}.\,u(t_0')\,v'(t_0'')} \tag{5}$$

Here

$$t_0'=b_1^{1/3}\,.\,\frac{\omega^2}{c^2}\left(\frac{4\pi e^2N_0}{m\omega^2}-1\right)$$

$$t_0''=b_2^{1/3}\,.\,\frac{\omega^2}{c^2}\left(\frac{4\pi e^2N_0}{m\omega^2}-1\right)$$

In expression (4), the second component of the right part is a complex amplitude of the reflected wave on plane $Z=Z_1$. Depending on the frequency of the falling wave, arguments t_0' and t_0'' for the assumed value of electron concentration N_0 on plane $Z=0$ may have both positive and negative values, or be reduced to zero. In all cases, expression (5) remains true; therefore, it is rather convenient for calculations, and we shall use it for the purpose of investigating the E-region of the ionosphere in conformity with modern concepts of monotonic density rise in this region. It follows immediately from equation (4) that the module of the coefficient of reflection from the ionized medium equals unity, which is evident for the case considered because the medium is assumed to be nonabsorbing.

Now, let us turn to the triangular law of ionization change when an effective electron density has a maximum ($N_0=N_{max}$ on plane $Z=0$). In this case the electric field strength at the boundary $(Z=Z_1)$ of the ionized medium will be determined by the following expression

$$E=1+M\,.\,\exp\left(-j\frac{m\omega^2}{3\pi ce^2K_1}\right).\exp\left(j\frac{\pi}{2}\right) \tag{6}$$

for $Z=Z_1$ where

$$M=\frac{\tau-q+j\,.\,(n+p)}{\tau+q+j\,.\,(n-p)}$$ and

$$\tau=u'(t_0')\,v(t_0'')+\left(\frac{K_2}{K_1}\right)^{1/3}u(t_0')\,v'(t_0'')$$

$$n=u'(t_0')\,u(t_0'')+\left(\frac{K_2}{K_1}\right)^{1/3}u(t_0')\,u'(t_0'')$$

$$p=v'(t_0')\,v(t_0'')+\left(\frac{K_2}{K_1}\right)^{1/3}v(t_0')\,v'(t_0'')$$

$$q=v'(t_0')\,u(t_0'')+\left(\frac{K_2}{K_1}\right)^{1/3}v(t_0')\,u'(t_0'') \tag{7}$$

Airy's functions and their first derivatives entering expressions (5) and (7) are tabulated, and therefore numerical calculations of equations (4) and (6) present no difficulty. Expression (6) will be used in studying reflection from the F-region of the ionosphere.

Let us consider particular cases of (4) and (6). We shall begin with monotonic ionization distribution.

(1) Let the effective electron concentration N_0 (on plane $Z=0$) to be so high that for the

given frequency of the falling wave, arguments t_0' and t_0'' in (5) are positive and have a great value ($t_0' \gg 1$ and $t_0'' \gg 1$). Then it is possible to make use of asymptotic formulas for Airy's functions, and expression (4) will have a rather simple form:

$$E = 1 + \exp\left(-j\frac{m\omega^3}{3\pi c \cdot e^2 \cdot K_1}\right) \cdot \exp\left(j\frac{\pi}{2}\right) \quad (8)$$

for $Z = Z_1$. Expression (8) shows that ionization over the level $Z = 0$ does not exercise influence over the reflected wave.

(2) Electron concentration N_0 is so low that for the assumed frequency arguments, t_0' and t_0'', are negative and high by absolute values [$|t_0'| \gg 1$ and $|t_0''| \gg 1$].

In this case (4) is transformed as follows:

$$E = 1 + \exp\left\{-j\frac{m\omega^3}{3\pi ce^2}\left[\left(\frac{1}{K_2} - \frac{1}{K_1}\right) \times \right.\right.$$

$$\left.\left. \times \left(1 - \frac{4\pi e^2 N_0}{m\omega^2}\right)^{3/2} + \frac{1}{K_1}\right]\right\} \cdot \exp\left(j\frac{\pi}{2}\right) \quad (9)$$

for $Z = Z_1$. Expression (9) is simplified on the condition that

$$\frac{4\pi e^2 N_0}{m\omega^2} \ll 1$$

and assumes, for $Z = Z_1$ the following form:

$$E = 1 + \exp\left\{-j\left[\frac{m\omega^3}{3\pi ce^2 K_2} - \right.\right.$$

$$\left.\left. - 2\frac{\omega}{c} \cdot N_0\left(\frac{1}{K_2} - \frac{1}{K_1}\right)\right]\right\} \cdot \exp\left(j\frac{\pi}{2}\right) \quad (10)$$

If we proceed from the laws of geometric optics and determine the phase difference $\phi(\omega)$ between the falling and reflected waves, it is possible to obtain expressions (8), (9) and (10), with the exception of factor $\exp(j\pi/2)$ which is obtained only in the case of a strict consideration of the problem.

For instance, for the derivation of equation (9), the phase difference is evidently determined by the sum of the following integrals:

$$\phi(\omega) = 2\frac{\omega}{c}\left[\int_{z_1}^{0} n_1(Z)\,dZ + \int_{0}^{z_0} n_2(Z)\,dZ\right]$$

where indexes of refraction $n_i(Z) = \sqrt{[\varepsilon_i(Z)]} = \sqrt{\left[1 - \frac{4\pi e^2}{m\omega^2}(N_0 + k_i Z)\right]}$ while $n_2(Z_0) = 0$.

Knowing the phase difference, $\phi(\omega)$, of the falling and reflected waves, one can easily determine a group signal delay time $\Delta t_{gr} = \phi'(\omega)$.

If the falling wave frequency is sufficiently low for the given electron concentration N_0, at the level $Z = 0$, and, consequently (8) is valid, the group delay time will be

$$\Delta t_{gr} = \frac{m\omega^2}{\pi ce^2 K_1} \quad (11)$$

On the contrary, for the case of high frequency when expression (9) is true, one obtains

$$\Delta t_{gr} = \frac{m\omega^2}{\pi ce^2}\left\{\left(\frac{1}{K_2} - \frac{1}{K_1}\right)[\varepsilon(0)]^{3/2} + \frac{1}{K_1}\right\} +$$

$$+ \frac{4N_0}{c}\left(\frac{1}{K_2} - \frac{1}{K_1}\right)[\varepsilon(0)]^{\frac{1}{2}};\left[\varepsilon(0) = 1 - \frac{4\pi N_0 e^2}{m\omega^2}\right] \quad (12)$$

or, when $\frac{4\pi N_0 e^2}{m\omega^2} \ll 1$,

$$\Delta t_{gr} = \frac{m\omega^2}{\pi ce^2 k_2} + 2 \cdot \frac{N_0}{c} \cdot \left(\frac{1}{K_1} - \frac{1}{K_2}\right) \quad (13)$$

In the case of the considered structure of region E with monotonic change of electron concentration, the meaning of the critical frequency concept changes. The critical frequency ω_{cr} is in fact the frequency ω_0 for which the dielectric permeability is reduced to zero at the level $Z = 0$, where a change of ionization gradient takes place (from value K_1 to K_2):

$$\varepsilon(0) = 1 - \frac{4\pi N_0 e^2}{m\omega_0^2} = 0 \quad (14)$$

Bearing in mind that in ionosphere region E the ionization gradient $K_1 \gg K_2$, we shall steadily increase the frequency beginning with values considerably lower than ω_0. Then according to expressions (11) and (12) the group delay time and the rate of its change with frequency for $\omega < \omega_0$ will be much lower

than for $\omega > \omega_0$. Thus, at the transition over frequency ω_0, the effective altitude of the considered region of ionized medium will rapidly increase, which corresponds to the specific form of altitude-frequency characteristics of the ionosphere. For ionization distribution shown in Fig. 3 the altitude-frequency characteristic (Fig. 6) is plotted with the help of

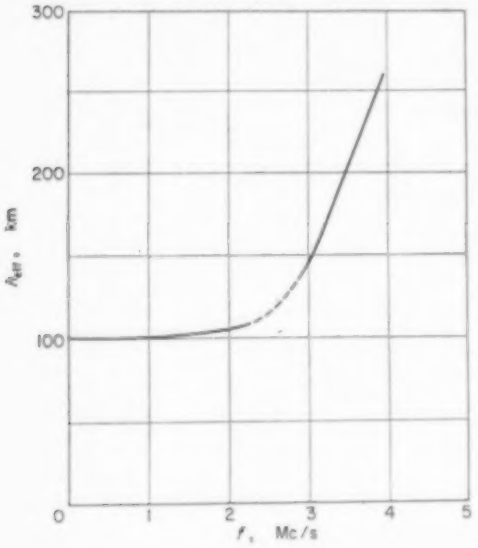


Fig. 6. The altitude–frequency characteristics of the E-region in the ionosphere.

expressions (11)–(12). It should be noted that validity of the formulas used is disturbed in the neighbourhood of frequency ω_0. This region is shown by a dotted section in Fig. 6.

Now let us turn to ionosphere region F which contains the maximum of electron concentration. Let us consider consequences which appear upon approximating electron distribution in this region by a triangular layer with the maximum at the level $Z=0$.

(3) If the maximum concentration N_{max} for a given frequency of the incident wave is so high that arguments $t_0' \gg 1$ and $t_0'' \gg 1$, then (6) is converted to (8). (In this case we use asymptotic formulas for Airy's functions.)

(4) For the critical frequency when arguments t_0' and t_0'' are reduced to zero, factor M in (6) is simplified:

$$M = \frac{1}{2} - j \cdot \frac{\sqrt{3}}{2} \cdot \frac{1 - \left(\frac{K_2}{K_1}\right)^{1/3}}{1 + \left(\frac{K_2}{K_1}\right)^{1/3}} \qquad (15)$$

For a symmetrical layer ($K_2 = K_1$) and $|M| = 0{\cdot}5$, i.e. in case of reflection, the amplitude decreases by 50 per cent. If we reduce gradient K_2 from value $K_2 = K_1$ to $K_2 = 0$, infiltration of energy through the ionized layer will decrease and cease when $K_2 = 0$ and ($|M| = 1$), i.e. when ionization after reaching the maximum does not already change with an increase of the altitude. Function $|M| = f(K_2/K_1)$ in a considerable part of range $0-1 \approx 0{\cdot}5$, and only when the argument K_2/K_1 approaches zero (rapidly rising), does it tend to approach unity.

(5) When the frequency is changed from values less than the critical one ($\omega < \omega_{cr}$) to values exceeding it ($\omega > \omega_{cr}$) the electric field amplitude of the reflected wave $|M|$ (entering expression (6)) will rapidly decrease for $\omega \approx \omega_{cr}$. The curve in Fig. 7 illustrates the decrease of the reflected field intensity at the transition over the critical frequency.

Up to now we have neglected absorption in radiowave propagation through an ionized medium. Let us consider the effect of absorption on a reflected wave.

Absorption of radio waves

For the studied laws of altitude-distribution of ionization it is possible to obtain a strict solution of the problem taking into account absorption, assuming that the number of collisions in formula (1) is a constant value. The general solution of the equation can be written in terms of Bessel functions of order one-third. For two regions of the ionized medium ($Z_1 \leqslant Z \leqslant 0$ and $Z > 0$) complex arguments of these functions for the case of monotonic rise of the ionization will be

$$\xi_i = \frac{2}{3} b_i^{1/2} \cdot \sigma_i^{3/2} \qquad (16)$$

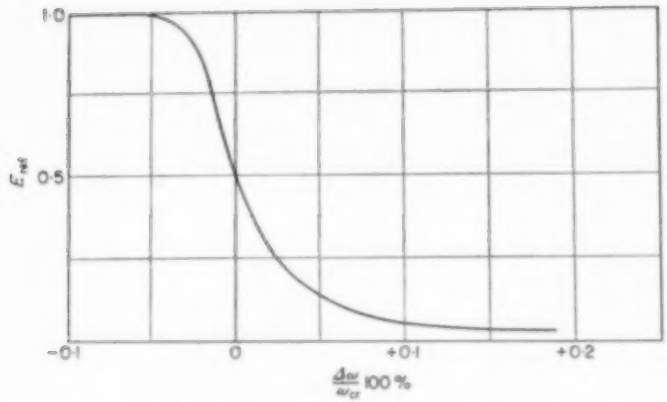


Fig. 7. The decrease of reflected field intensity at transition over the critical frequency.

while

$$b_i^{1/2} = \frac{mc^2}{4\pi e^2}\,\frac{1}{K_i}\left(1 - j\,\frac{V}{\omega}\right) \tag{17}$$

and

$$\tau_i = \frac{\omega^2}{c^2}\left[1 - \frac{4\pi e^2}{m\omega^2\left(1 - j\,\frac{V}{\omega}\right)}\,(N_0 + K_i \,.\, Z)\right] \tag{18}$$

where $i = 1, 2$. When Z is increased, vectors corresponding to complexes ξ_i rotate clockwise on the complex plane. For $Z \longrightarrow \infty$ the vector determined by complex ξ_2 will lie in the second quadrant and will be inclined at an angle of $\frac{1}{2}\chi$ to the imaginary axis while $\chi = \operatorname{arc\,tg} V/\omega$. As usual, for the ionosphere $V/\omega \ll 1$ and correspondingly, angle χ has a low value.

The solution is performed in such a way that for the limit $V \longrightarrow 0$ relation (4) is obtained.

First of all two specific cases should be noted.

(i) If for the pre-set frequency electron concentration N_0 is low—so that complexes $-t_0' = b_1^{1/3} \,.\, \tau_0$ and $-t_0'' = b_2^{1/3} \,.\, \tau_0$ [$\tau_0 = \tau$ for $Z = 0$] are located in the fourth quadrant of the complex plane—the solution can be conveniently expressed in functions $u(t) = u_-(t)$, $v(t) = v_-(t)$. ($u_-(t)$ and $v_-(t)$ are related with Bessel functions in the same manner as the Airy functions of the real negative argument, (see the Appendix).

(ii) For the case of high concentration of N_0 when complexes $-t_0'$ and $-t_0''$ are located in the third quadrant of the complex plane, the solution can be conveniently expressed in functions $u(t) = u_+(t)$, $v(t) = v_+(t)$. ($u_+(t)$ and $v_+(t)$ are related with Bessel functions in the same manner as the Airy functions of the real positive argument). The values of Airy complex argument functions $u(t)$ and $v(t)$ can be directly obtained from the "Tables of the modified Hankel functions of order one-third and of their derivatives."[6]

The electric field strength at the boundary of the ionized medium for monotonous change of ionization will be determined by the following expression

$$E = 1 + \frac{\tau + j \,.\, p}{\tau - jp}\exp\left(-\frac{m \,.\, v \,.\, \omega^2}{3\pi c e^2 K_1}\right) \times$$
$$\times \exp\left(-j\,\frac{m\omega^3}{3\pi c e^2 K_1}\right) \,.\, \exp\left(j\,\frac{\pi}{2}\right) \tag{19}$$

while complexes

$$\tau = u'(t_0') \,.\, v(t_0'') - \left(\frac{K_2}{K_1}\right)^{1/3} \,.\, u(t_0') \,.\, v'(t_0'')$$
$$p = v'(t_0') \,.\, v(t_0'') - \left(\frac{K_2}{K_1}\right)^{1/3} \,.\, v(t_0') \,.\, v'(t_0'') \tag{20}$$

where $t_0' = -b_1^{1/3}\,\tau_0$ and $t_0'' = -b_2^{1/3}\,\tau_0$;

$$\tau_0 = \frac{\omega^2}{c^2}\left[1 - \frac{4\pi e^2 N_0}{m\omega^2\left(1 - j\frac{V}{\omega}\right)}\right].$$

If in the right part of equation (19) we pass to the limit for $V \to 0$, we shall obtain expression (4) derived earlier, which is true for a non-absorbing medium. For the case of high concentration N_0 when arguments $-t_0'$ and $-t_0''$ are located in the third quadrant of the complex plane, and besides this $|t_0'| \gg 1$ and $|t_0''| \gg 1$:

$$E = 1 + \exp\left(-j\frac{m\omega^2}{3\pi c e^2 K_1}\right) \times$$

$$\times \exp(-\Gamma) \cdot \exp\left(j\frac{\pi}{2}\right) \qquad (21)$$

for $Z = Z_1$

where the absorption coefficient of reflected wave is

$$\Gamma = \frac{V \cdot m\omega^2}{3\pi c e^2 K_1} \qquad (22)$$

For the case of low concentration of N_0, when arguments $-t_0'$ and $-t_0''$ occur to be in the fourth quadrant and $|t_0'| \gg 1$ and $|t_0''| \gg 1$, we shall obtain from (20) on the condition that $\left(\frac{V}{\omega}\right)^2 \ll 1$ an expression identical to (9) but containing factor $\exp(-\Gamma)$ where

$$\Gamma = \frac{V\omega^2 m}{3\pi c e^2}\left[\left(\frac{1}{K_2} - \frac{1}{K_1}\right)[\varepsilon(0)]^{3/2} + \frac{1}{K_1}\right] +$$

$$+ 2\frac{V}{c} N_0 \left(\frac{1}{K_2} - \frac{1}{K_1}\right)[\varepsilon(0)]^{1/2} \qquad (23)$$

where $\varepsilon(0) = 1 - \frac{4\pi e^2 N_0}{m\omega^2}$

It should be noted that formula (23) can be derived on the basis of geometrical optics where

$$\Gamma = 2 \cdot \frac{\omega}{c}\int_{z_1}^{0} \kappa_1\, dz + 2 \cdot \frac{\omega}{c}\int_{0}^{z_0} \kappa_2 dz$$

while $\kappa_i = \frac{2\pi e^2 V}{m\omega^3 n_i}\,(N_0 + K_i Z)$; $n_i(Z_0) = 0 (i = 1, 2)$.

If condition $4\pi e^2 N_0 / m\omega^2 \ll 1$ is fulfilled, we shall find from (23) that

$$\Gamma = \frac{V \cdot m \cdot \omega^2}{3\pi c e^2 K_2} \qquad (24)$$

For the ionospheric E-region $K_2 < K_1$. With a rise in frequency, as follows from (22) and (24), absorption increases first with a lower rate, and then with the greater one. We have seen earlier that at the transition over frequency ω_0 the effective altitude of the considered region of ionized medium sharply rises, and at the same time the reflected wave absorption will show rapid increase which, as is known, is proved by experiment. Assuming effective number of collisions $V = 10^5$ it is possible to calculate the absorption coefficients for ionization distribution shown in Fig. 3.

$\Gamma = 1{\cdot}24$ for frequency $\omega = 10^7$

$\Gamma = 21{\cdot}5$ for frequency $\omega = 2 \times 10^7$

In the latter case the reflected wave absorption appears to be rather high. If the decrease of the effective number of collisions with altitude is taken into account, for instance following the exponential law, the values of reflection coefficients will correspondingly have lower magnitude.

Now let us pass to absorption in ionosphere region F. In approximating ionization distribution by the triangular layer the following expression for the electric field strength at the lower boundary of ionized medium is obtained:

$$E = 1 + M \cdot \exp\left(-j\frac{m\omega^2}{3\pi c e^2 K_1}\right) \cdot \exp\left(-\frac{Vm\omega^2}{3\pi c e^2 K_1}\right) \times$$

$$\times \exp\left(j\frac{\pi}{2}\right) \text{ for } Z = Z_1 \qquad (25)$$

where $M = \frac{\tau - q + j(n+p)}{\tau + q + j(n-p)}$

By its composition (25) is identical to (6) obtained earlier. Therefore it is sufficient to

write the expression only for one complex parameter, for example, τ:

$$\tau = u'(t_0')\,v(t_0'') + \left(\frac{K_2}{K_1}\right)^{1/3} \cdot u(t_0')\cdot v'(t_0'')$$

As before, reflection depends on the location of arguments $-t_0'$ and $-t_0''$ on the complex plane. In case of high ionization for a given frequency, we shall obtain the former expression (21). In the presence of absorption arguments t_0' and t_0'' are always different from zero. In case of low ratio $V/\omega \ll 1$, modules $|t_0'|$ and $|t_0''|$ will be low in the region of the critical frequency, and then (25) yields the following expression for the field strength

$$E = 1 + M\cdot\exp\left(-j\frac{m\omega^3}{3\pi ce^2K_1}\right)\cdot\exp\left(-\frac{Vm\omega^3}{3\pi ce^2K_1}\right)\times \times\exp\left(j\frac{\pi}{2}\right) \qquad (26)$$

Expression (26) shows that besides absorption determined by the value

$$\Gamma = \frac{mV\omega^2}{3\pi ce^2K_1}$$

the wave amplitude decreases which is determined by factor M (see expression (15)) as a result of energy infiltration through the ionized layer; the process of partial reflection of a wave from the layer takes place. The amplitude of the reflected wave sharply decreases if the frequency becomes higher than the critical one.

3. NONSTATIONARY PROCESSES AT THE REFLECTION OF AN ELECTROMAGNETIC SIGNAL FROM HETEROGENEOUS IONIZED LAYER

Let us suggest that an electromagnetic signal of the following form falls upon the interface of the ionized medium:

$$E_{\text{fall}} = \exp(j\omega_c t), \qquad \text{for } t>0$$
$$E_{\text{fall}} = 0 \qquad \text{for } t<0$$

(ω_c = carrier frequency).

Let us consider the reflected signal. For this purpose the electric field strength of the reflected wave will be presented in the form of a Fourier integral. If absorption is neglected, it is convenient for the approximate calculation of a Fourier integral to use the stationary phase method. It should be that here the classical Thomson formula will somewhat change, for we are compelled to consider the case in which the point of the stationary phase is located near the limits of integration. The solution of the given problem was described by the author in 1940.[7] To illustrate this method we shall assume that a phase shift between the incident and the reflected waves will be determined, for instance, by equation (8)

$$\phi(\omega) = \alpha\omega^2 - \frac{\pi}{2} \quad \text{while } \alpha = \frac{m}{3\pi ce^2K_1}$$

Then the envelope form of transient signal will be determined by relations

$$|E_{\text{ref}}| = \frac{1}{\sqrt{2}}\sqrt{\left[\frac{1}{2} - C(u)\right]^2 + \left[\frac{1}{2} - S(u)\right]^2}$$

for $t < 3\alpha\omega_c^2$

$$|E_{\text{ref}}| = \frac{1}{\sqrt{2}}\sqrt{\left[\frac{1}{2} + C(u)\right]^2 + \left[\frac{1}{2} + S(u)\right]^2}$$

for $t > 3\alpha\omega_c^2$ (27)

Here, $C(u)$ and $S(u)$ are Fresnel functions, where

$$u = \frac{1}{12\alpha\omega_c}\cdot\tau^2;\ \tau = t - 3\alpha\omega_c^2$$

The use of the super-position method and Duamel's theorem enable us to plot reflected impulses if the incident impulses have different forms. For ionization distribution in regions E (Fig. 3) and F (Fig. 4) the envelope curves of transient signals (Fig. 8) were plotted for two carrier frequencies of which one is $\omega_1 = 10^7$ and the other $\omega_2 = 6{\cdot}28\times10^7$. In the second case the delay of a signal appears to be greater, and the transient process takes place more slowly than in the first case. The group delay of a signal in the first case is $6{\cdot}37\times10^{-4}$ sec, and in the second case $-2{\cdot}2\times10^{-5}$ sec.

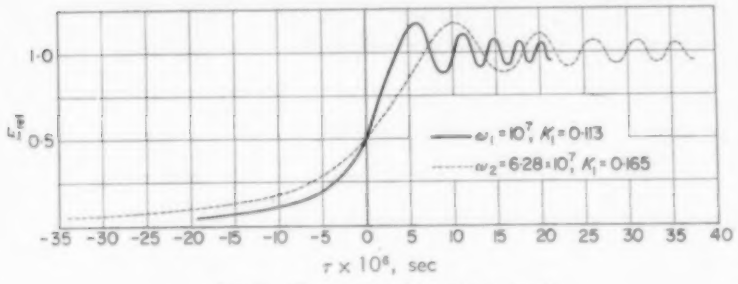


Fig. 8. The curves of transient signals.

At present our knowledge in regions of ionization in the high atmosphere is not sufficient and requires further correction. Besides, atmosphere ionization may undergo great changes. The curves for altitude-distribution of ionization density may have a lot of relative maxima and minima (e.g. see Fig. 2). The present work is based on some of the most simple types of altitude-distribution of ionization. With the accumulation of new experimental data, the conclusions contained in this work should naturally undergo certain alterations and corrections.

CONCLUSIONS

1. This work considers electromagnetic wave propagation in the ionosphere on the basis of wave treatment of the problem with ionization distribution corresponding to the latest experimental data. The solution expressed in Airy functions is convenient for investigation of the reflected field in the frequency range considered.

2. In conformity with new concepts ionosphere region E shows a sharp change of ionization gradient; however, a continuous and approximately monotonic rise in ionization with altitude is preserved. Therefore, with the rise of frequency the level of the electromagnetic wave reflection is steadily shifted to higher regions of the ionosphere. According to the former concepts of sharply limited layers in the ionosphere, at the transition over the critical frequency, the wave reflection region shifts skip-like from layer E to higher ionosphere layers F_1 or F_2. The critical frequency is now actually frequency ω_0 at which the dielectric permeability for ionization density N_0, corresponding to the level of sharp change in the ionization gradient, is reduced to zero.

3. At the transition over frequency ω_0 the effective ionosphere altitude sharply increases with simultaneous rise of the reflected wave absorption so that the received signal intensity becomes rather low.

4. The envelope curves of transient signals reflected from ionosphere regions E and F are plotted with ionization distributions corresponding to the latest experimental data.

5. For low absorption in the ionized medium the results obtained by the strict solution of the problem are approximately reconciled with conclusions of geometric optics.

Acknowledgements—I consider it to be my pleasant duty to express my acknowledgements to V. I. Balashova and T. S. Romanova for their help in carrying out this work.

APPENDIX

Relation of Airy and Bessel functions.

(1) Airy function of the negative argument

$$u(-t) = \sqrt{\left(\frac{\pi}{3}t\right)}\,[J_{-1/3}(x) - J_{1/3}(x)];$$

$$v(-t) = \tfrac{1}{2}\sqrt{(\pi t)}\,[J_{-1/3}(x) + J_{1/3}(x)] \text{ where } x = \tfrac{2}{3}t^{3/2},\ t > 0.$$

(2) Airy functions of the positive argument

$$u(t) = \sqrt{\left(\frac{\pi}{3}t\right)}\,t\,[I_{-1/3}(x) + I_{1/3}(x)];$$

$$v(t) = \tfrac{1}{2}\sqrt{(\pi t)}\,[I_{-1/3}(x) - I_{1/3}(x)] = \frac{1}{\sqrt{(3\pi)}}\sqrt{t}\,.\,K_{1/3}(x).$$

Asymptotic expressions for Airy functions are the following if we limit ourselves only to the first member of the expansion:

$$u(-t)=t^{-1/4}\cdot\cos\left(x+\frac{\pi}{4}\right);$$

$$v(-t)=t^{-1/4}\cdot\sin\left(x+\frac{\pi}{4}\right);$$

$$u(t)=t^{-1/4}\cdot\exp x,$$

$$v(t)=\tfrac{1}{2}t^{-1/4}\cdot\exp -x.$$

(3) For the complex argument

$$u_-(t)=\left(\frac{\pi}{3}t'\right)^{1/2}\cdot[J_{-1/3}(x)-J_{1/3}(x)];$$

$$v_-(t)=\tfrac{1}{2}(\pi t)^{1/2}\cdot[J_{-1/3}(x)+J_{1/3}(x)];$$

where $x=\frac{2}{3}\cdot t'^{3/2}$, $-t=-t'$.

$$u_+(t)=\left(\frac{\pi}{3}t\right)^{1/2}\cdot[I_{-1/3}(x)+I_{1/3}(x)];$$

$$v_+(t)=\tfrac{1}{2}(\pi t)^{1/2}\cdot[I_{-1/3}(x)-I_{1/3}(x)];$$

where $x=\frac{2}{3}t^{3/2}$.

REFERENCES

1. K. I. GZINGAUZ, *Dokl. Akad. Nauk SSSR* **120,** 1234 (1958).
2. J. SEDDON, A. PICKAR and J. JACKSON, *J. Geophys. Res.* **59,** 513 (1954).
3. J. JACKSON, *J. Geophys. Res.* **61,** 107 (1956).
4. G. BURGESS, *Frontier to Space*. Chapman and Hall, London (1955).
5. V. A. FOK, *Tables of Airy's Functions*, Moscow (1946).
6. *Tables of the Modified Hankel Functions of Order One-Third and of Their Derivatives*. Cambridge, Mass., 1945.
7. L. ZHEKULIN, *Bull. Akad. Nauk SSSR, ser. fiz.* **4,** 409 (1940).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 121–129. Printed in Great Britain.

METEORITIC DUST MEASURED FROM EXPLORER I

MAURICE DUBIN*
Air Force Cambridge Research Center

(*Received* 29 *July* 1959)

Abstract—The impacts of micrometeorites upon the cylindrical shell of the satellite Alpha 1958 were monitored by a calibrated piezo-electric detector. Results of these measurements are presented, including a description of the instrumentation, the calibration procedure, and the amount and type of data recorded at the ground stations. The total data sample over a period of 12 days of telemetering consisted of 78,890 sec of telemetered data with an excellent signal to noise ratio. During this time 153 signals denoting impacts by micrometeorites were recorded. This is equivalent to $1{\cdot}7\times10^{-2}$ impacts/m^2/sec on the earth. The calibration indicates that for a mean impact velocity of 30 km/sec the particles striking the satellite had a mass of 8×10^{-10} g or larger, if no large deviations in momentum transfer occur during the formation of hypervelocity craters. Thus, the accretion rate of extraterrestrial dust upon the earth may be estimated at 10,000 tons/day during the month of February 1958. A variation from day to day of the influx rate as large as an order of magnitude is evident from the data. A "shower" of dust particles has been found for the third day in orbit. A variation of the influx of particles may also be attributed to a diurnal effect from the earth's rotation and its heliocentric velocity.

1. INTRODUCTION

Meteoritic dust, impacting upon the cylindrical shell of satellite Alpha 1958, Explorer I, was monitored with a detector incorporating a piezo-electric transducer and a transistorized amplifier. These measurements represent the first substantial, direct measurement of interplanetary matter in the vicinity of the earth.

Measurements of extra-terrestrial dust, prior to the use of direct probing with rockets and satellites, were carried out by ground-based methods, optical and radar observations, and measurements of the corona and zodiacal light. The development of rockets and their subsequent use for research in the outer atmosphere, and more recently the success in launching satellites,[1, 2] have added a new dimension for direct measurements of extra - terrestrial material.

From previous ground-based investigations, a large amount of information about interplanetary matter was available for defining the design specifications of a direct probing experiment. Above the atmosphere, beginning at the *E*-region, the lower limit of the geocentric velocity of the cosmic dust is the escape velocity from the earth, 11 km/sec. Because the cosmic dust moves in heliocentric orbits, the upper limit of the geocentric velocity is close to 70 km/sec. The problem of initial interest, using the new method of direct probing of these hypervelocity particles from rockets and satellites, was the measurement of cosmic dust beyond the detection limit of optical and radio measurements, approximately visual magnitude 10, for masses less than 10^{-4}g.[3]

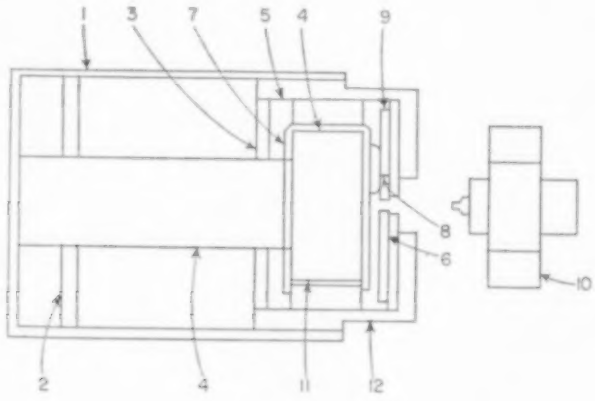
2. INSTRUMENTATION

As reported,[2, 4, 5] the measurements of micrometeorites were made using a crystal transducer and transistorized amplifier to detect the impulse of the impacting meteoroid from a satellite. This type of instrumentation is reliable, compact and, although of simple design, is sufficiently sensitive to provide a good data sample during the time that the experiment was in operation.

* Present address: National Aeronautics and Space Administration, Washington, D.C.

Three reasons for using this type of detector are: (1) the area of the sensitive surface exposed to meteoritic impacts could be made as large as feasible using a single sensing element; (2) the high sensitivity of the detector permitted the detection of a large data sample since the number density of particles in interplanetary space increases with decreasing mass; and (3) it had been found by calibration that the detection system was sensitive to the momenta of the impacting particles, and therefore the mass of the impacting particles could be determined with an error depending upon the first power of the velocity.

The sensing element was a lead zirconate crystal mounted in a machined cylindrical housing in contact with the cylindrical shell of the satellite. The transducer was resonant-tuned during construction to give a maximum signal output at ultrasonic frequencies around 100 kc/s. A diagram of the crystal microphone is shown in Fig. 1. The pulse from the crystal was amplified by a transformer-coupled, transistorized amplifier with a gain of 90 dB (Fig. 2). The amplifier was also tuned to 100 kc/s $\pm$ 5 kc/s by slugs in the transformers. The impact of a micrometeorite on the satellite skin causes an exponentially damped ringing in the material of the shell. The 100 kc/s component of this impulse observed at the output of the amplifier has a rise time less than a microsecond, and decreases by an order of magnitude in amplitude after a few oscillations. This impulse was rectified at the last stage of the amplifier to give a direct current pulse used to switch a storage circuit, consisting of an Eccles–Jordan circuit (Fig. 3). This bistable multivibrator modulated one of the four subcarrier oscillators of the high powered transmitter on the satellite. A change in position of the bistable multivibrator shifted the subcarrier oscillator frequency approximately 100 cycles about a center frequency of 960 c/s. Thus, the information rate required was dependent upon the rate of impacts by micrometeorites. The micrometeorite information was carried on one of four channels of telemetry on the high power transmitter of Explorer I.



1 Mic. aluminum case
2 Plastic ring support No. 1
3 Plastic ring support No. 2
4 Brush crystal PZT–5
5 Rubber spacer
6 Copper plate
7 Copper foil
8 Solder bead
9 Plastic disc
10 Microdot connector
11 Foil slot
12 Mic. end cap

Fig. 1. Micrometeorite detector microphone.

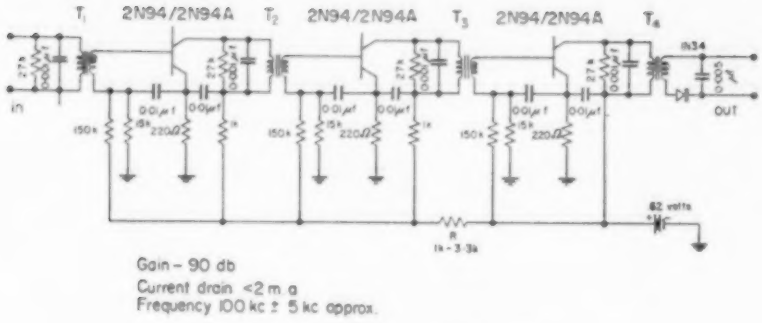


Fig. 2. Amplifier for meteoritic detection.

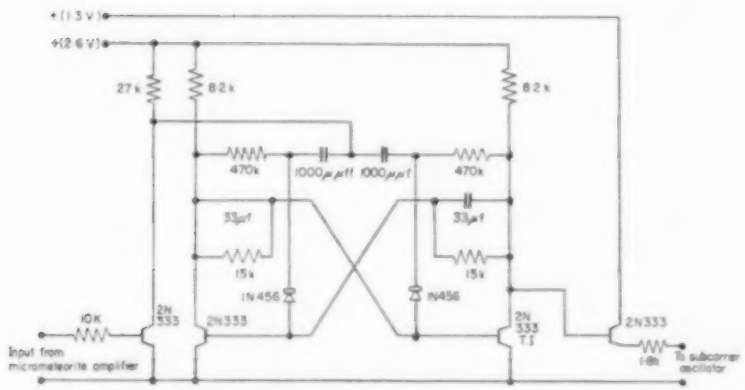


Fig. 3. Pulse storage circuit.

This transmitter radiated 60 mW of power on a frequency of 108·03 Mc/s for the first 12 days of the satellite orbit. Fig. 4 is a photograph of the micrometeorite detector, the bistable multivibrator, and the mercury batteries used to supply three weeks of power. Fig. 5 is a photograph of the satellite instrumentation. The microphone was mounted on the ribbed support. Impact information was fed to the high power transmitter mounted at the base of the payload.

3. CALIBRATION

The micrometeorite detection equipment was calibrated by impacting particles on the satellite shell. In this first, elementary satellite experiment, particle impacts above a fixed threshold level could be monitored. The calibration was carried out to determine the effective threshold level for this equipment.

In order to calibrate the equipment it was necessary to determine the characteristic parameters of the particle impact. The particle is characterized by its mass and the velocity of impact. A series of experiments was carried out to determine the sensitivity of particle impacts by impact tests over a large range of mass and velocity. Theoretically, this type of detector is sensitive to the momentum of the impact. It was found that the momentum dependence was valid over the entire range of mass and velocity producible by laboratory methods for the range of sensitivity of the micrometeorite detection equipment.

The low ranges for velocity of impact were obtained by dropping glass spheres on the

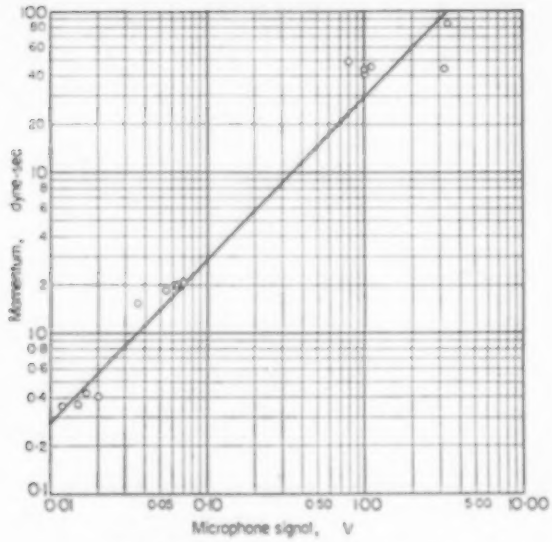


Fig. 6. Microphone calibration curve.

sensitive surface. In this manner a velocity range of 20 cm/sec to 400 cm/sec was obtained using spheres with diameters ranging from 65 μ to 500 μ (0·4–200 μg). These tests were corrected for viscous drag, the dropping height varying from a couple of millimeters to a few centimeters. Since these impacts were elastic, the impulses were measured in some cases through several rebounds. This calibration indicated that the equipment was sensitive to the mometum of impact.

The calibration at the high range of velocity was carried out with cylindrical, explosive-charge accelerators[6]. Single steel particles with diameters from 38 to 300 μ were accelerated to velocities as high as 4 km/sec and impacted upon the sensitive surface. The velocity and impulse were measured. Fig. 6 is a curve of the signal obtained from the microphone as a function of the momentum of the impact. These tests indicated that the micrometeorite detection equipment was sensitive to the momentum of the impact over a range of momenta covering four orders of magnitude. At velocities of 4 km/sec, the impact is inelastic and a hypervelocity crater is formed on the impacted surface. Comparisons were made on the same equipment by dropping glass spheres. After corrections were made for elastic and inelastic impacts, it was found that both methods gave consistent results within a factor of 2. Thus, the micrometeorite detection equipment was found to be sensitive to the momentum of the impact over a large range of masses and for a range of velocities from 2×10^{-4} km/sec to 4 km/sec.

The velocities of meteoroid impacts upon the satellite may vary between 10 km/sec and 70 km/sec. The momentum transferred to the impacting surface from the micrometeorite impact is measured with the acoustical detection system. At meteoroid velocities, no direct calibration measurements of the momentum transfer are available. There is a possibility that during crater formation the amount of surface material and meteoritic material exploded from the surface may result in a more efficient momentum transfer than obtained from simple momentum transfer. However, for calibrations performed at velocities of 4 km/s with crater formation this was not borne out. Although material was ejected during crater formation, the momentum transfer behaved like an inelastic collision. An adequately proven theory

Fig. 4. Micrometeorite detector for Explorer I.

Facing page 124

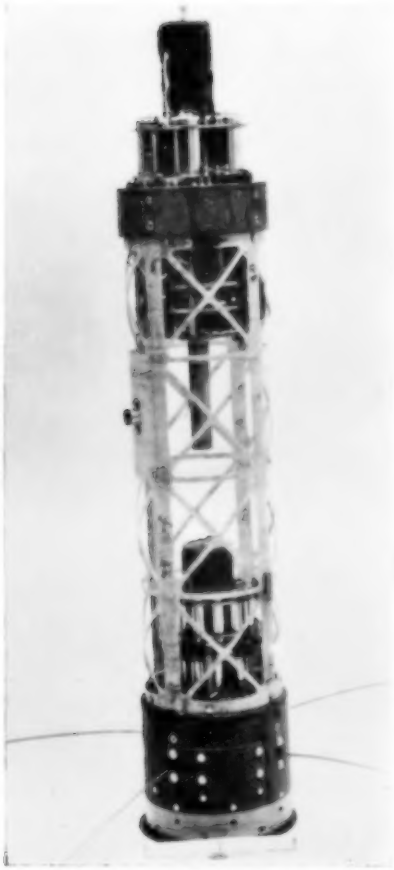


Fig. 5. Instrumented payload section of Explorer I.

of crater formation giving the mometum transfer at meteoritic velocities is not presently available.

The material ejected during crater formation will, of course, add to the momentum transferred. It appears, however, that at meteoritic velocities the momentum transfer with cratering is not much different from momentum transferred during elastic impacts. In any event, the error from the effects of cratering is probably no greater than the errors in the measurement from not knowing the actual velocity of impact and from the variation in sensitivity over the cylinder of the satellite. The overall error in determining the mass of micrometeorites impacting upon the satellite is less than a factor of 10.

The calibration of the flight model of the satellite equipment was made by dropping glass spheres. Fig. 7 gives the threshold momentum for elastic impacts required to shift the bistable multivibrator at a few of the points used in mapping the cylindrical shell. Twenty impacts were used to determine the values for each test point. Point *A*, directly over the microphone, had a threshold value of $1{\cdot}7 \times 10^{-3}$ g cm/sec. Point *D*, the least sensitive point on the cylinder, was diametrically opposite point *A* and at the far end of the cylinder. The impulse measure at *D* was $5{\cdot}2 \times 10^{-3}$ g cm/sec, only one-third the sensitivity of point *A*. Both ends of the cylinder were connected to insulators used in making the transmitting antennae. The insulating material transmitted impacts very poorly. The conical tip was 100 times less sensitive than point *A*, and the last stage motor was so insensitive to impacts that a value could not be determined. Thus, the sensitive area was a cylinder with diameter of 15·2 cm and length of 45·7 cm, an area of 0·23 m². The effective threshold momentum for this area was $2{\cdot}5 \times 10^{-3}$ g cm/sec.

4. RESULTS AND DISCUSSION

Satellite Alpha 1958 was successfully launched at 0348 U.T. on 1 February 1958 from Cape Canaveral, Florida. The orbit had an inclination angle of 33·4°, perigee altitude of 374 km, apogee at 2550 km, and period of 115 min. Data was received only while the satellite was over a receiving station, since there was no data storage system. Good signals were received on a single pass for as long as 10 min when the satellite was near apogee. Data from the micrometeorite experiment was recorded by regular telemetry stations and by Minitrack stations at Blossom Point, Maryland; Fort Stewart, Georgia; Havana, Cuba; Woomera, Australia; Quito, Equador; Lima, Peru; Antofagasta, Chile; Santiago, Chile; An-

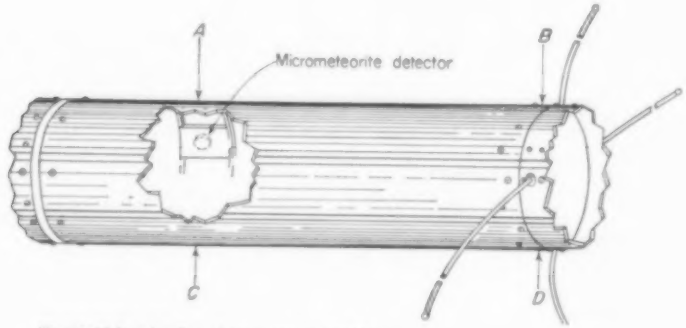


Fig. 7. Micrometeorite impact calibration on cylindrical shell of Explorer I.

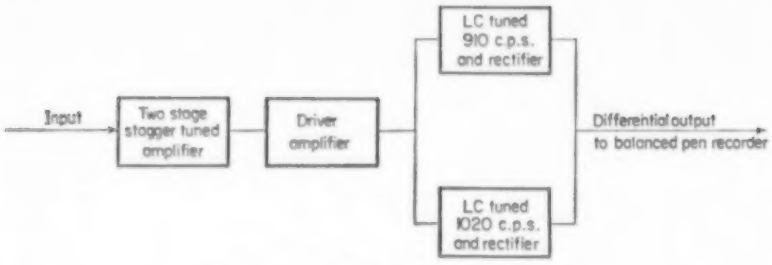

Input
Two stage stagger tuned amplifier
Driver amplifier
LC tuned 910 c.p.s. and rectifier
LC tuned 1020 c.p.s. and rectifier
Differential output to balanced pen recorder

Fig. 8. Discriminator for data reduction.

tigua, British West Indies, and San Diego, California. Eight Microlock stations in Florida, Singapore, Nigeria, California, etc., recorded each close passage of the satellite. Amateur radio operators, particularly in Kansas, North Carolina, Louisiana, Minnesota and Texas, also received good recordings.

The tapes were fed through a discriminator shown in block diagram form in Fig. 8. The discriminator had two tuned channels at 910 c/s and 1020 c/s to correspond to the high and low frequency positions of the subcarrier oscillator; the band-width was 25 cycles and the time constant was about 3 sec. The output of the discriminator was recorded on a two channel paper record, the second channel being used for the time code. The channel signal was also observed on multiplexed oscilloscopes, each having oscillator tuned horizontal sweeps. A Lisajou figure would form on one oscilloscope and switch to the other scope when the subcarrier oscillator changed frequency. A loudspeaker was also used to monitor by ear the characteristic tone of the signal. Fading of the signal could easily be recognized. The three methods of analysis were carried out simultaneously for nearly 400 rolls of 30-min tape, representing the amount of information received during the first 12 days of Alpha 1958. In this manner the signal could be classified into types *A*, *B* and *C*. Class *A* represented a signal with an excellent signal to noise ratio and was quite easily read. Class *B* was readable with some difficulty, and Class *C* was noisy and questionable. Fig. 9 is a section of Class *A* data recorded on paper and indicates the change observed during a switch of the subcarrier oscillator.

In Table 1, the results of the *A* and *B* data are given for each day. Both *A* and *B* telemetering signal times were noted for each station, and the number of switches of the subcarrier oscillator or hits is recorded. There were 153 hits recorded in the total class-*A*

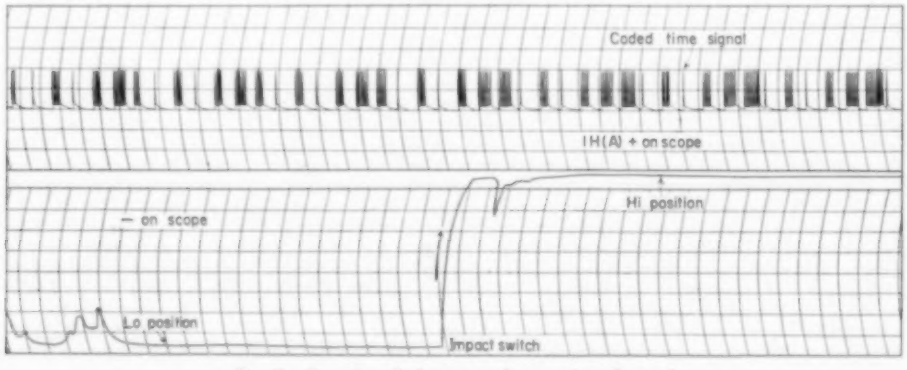


Fig. 9. Sample of *A*-type telemetering Record.

Table. 1. Compiled Data for Explorer 1—Daily Totals (Greenwich Date)

Date Feb.	*A* time (sec)	*A* hits	*A* rate $\times 10^3$ sec	*B* time (sec)	*B* hits	*B* rate $\times 10^3$	*A*+*B* time	*A*+*B* hits	*A*+*B* rate	Unseen hits
1	3905	15	3·84	421	—	—	4326	15	3·47	4
2	8584	39	4·54	580	1	1·72	9164	40	4·36	9
3	8634	57	6·60	1220	4	3·28	9854	61	6·19	—
4	7882	20	2·54	127	2	15·7	8009	22	2·75	7
5	6584	—	—	550	—	—	7134	—	—	8
6	5420	1	0·18	195	2	10·3	5615	3	0·53	6
7	7461	6	0·80	567	—	—	8028	6	0·75	8
8	6060	—	—	473	—	—	6533	—	—	6
9	7652	1	0·13	426	1	2·35	8078	2	0·25	6
10	7565	7	0·93	77	—	—	7642	7	0·92	6
11	6939	7	1·01	425	—	—	7364	7	0·95	13
12	2209	—	—	247	2	8·10	2456	2	0·81	1
Total	78895	153	1·94	5308	12	2·26	84203	165	1·96	84

time of 78,895 sec. This represents nearly 8 per cent of the total time of telemetering with the high power transmitter. There were only 12 hits recorded during the 5308 sec of *B* time. The rate of class-*A* hits, $1{\cdot}94 \times 10^{-3}$ hits/sec, compares well with the *B* rate of $2{\cdot}3 \times 10^{-3}$ hits/sec.

Figure 10 is a plot of the number of hits recorded during each Greenwich day. A large fluctuation from day to day is noticeable. On 5 February and 8 February no hits were recorded, although eight and six "unseen" hits were found (Table 1), an indication that the equipment was operating satisfactorily. "Unseen" represents information that an odd number of hits, and at least one hit, occurred between reception of the signal at two sequential receiving stations.

Although the data sample is small, an idea of the significance of the data may be obtained from the number of hits or the rate of hits as a function of local time. Fig. 11 is a plot of the number of hits in each hour of local time for the first 88 hits in the data sample. A diurnal dependence, resulting by the earth's heliocentric velocity, is evidenced, in that nearly ninety per cent of the 72 hits shown by the black bars occurred on the dawn side of the

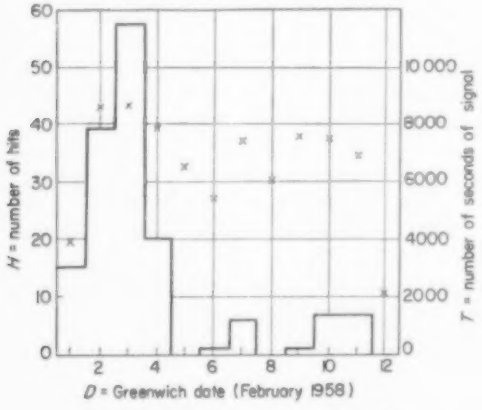


Fig. 10. Number of impacts per day.

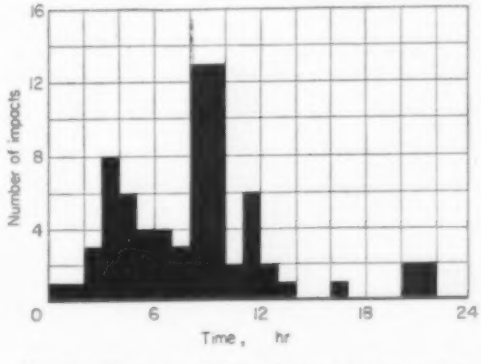


Fig. 11. Diurnal variation of meteoritic impacts on Alpha 1958.

earth between the hours of midnight and twelve noon. The three dashed lines at 16, 18 and 20 hr represent 16 hits and are probably a micrometeor "shower". All of these impacts were observed during a 4-hr interval on 3 February at two stations in the United States which were 2000 km apart, while the satellite was near perigee. The rate of impacts was about 50 times greater than the average rate of $1{\cdot}94 \times 10^{-3}$ hits/sec, and at a time of the day when the diurnal variation would normally give a minimum rate of influx of interplanetary material.

The impact rate of micrometeorites on Alpha 1958 was $1{\cdot}94 \times 10^{-3}$ sec^{-1}. Taking into account the sensitive area of 0·23 m^2, the rate was $8{\cdot}4 \times 10^{-3}$ impacts m^{-2} sec^{-1}. Because the earth is shielding the satellite, the impact rate for the same mass component of micrometeorites at one astronomical unit is nearly twice this value, or $1{\cdot}7 \times 10^{-2}$ impact m^{-2} sec^{-1}, on the upper surface of the satellite.

Based upon the calibration of the effective momentum threshold of $2{\cdot}5 \times 10^{-3}$ g cm/sec, assuming a linear momentum dependence is valid at meteor velocities ,and using a mean velocity of impact of 30 km/sec, the mass of the impacting particle may be computed. This mass is 8×10^{-10} g or greater at a velocity of 30 km/sec. The mass influx upon the earth would then be $1{\cdot}4 \times 10^{-11}$ g m^{-2} sec^{-1} for this component of interplanetary matter. This is equivalent to 6×10^5 g/day or 6×10^2 tons/day. If one may assume that a distribution of the mass and number of particles follows the pattern measured by Watson[7], then the total influx of meteoritic material from tenth visual magnitude to thirtieth visual magnitude would be approximately 10^4 tons/day over the earth's surface.

This value of 10^4 tons of interplanetary material falling upon the earth each day is a tentative one because of the assumptions upon which it is based. It was assumed that the distribution of cosmic dust was not much different from that given by Watson; this distribution may be measured directly from satellites. It was also assumed that since the micrometeorite detector on Alpha 1958 was sensitive to the momenta of the impacts to velocities as high as 4 km/sec, the momentum transfer remained proportional to the first power of the velocity of impact at meteor velocities. If this is not basically correct, the error may become large at the high end of the velocity distribution. The velocity of the impact must be measured also in future satellite experiments. It was further assumed that all impulses measured were meteoritic impacts. During the powered flight, the detector indicated a number of pulses, particularly while the motors of the last three stages were burning. The satellite was internally free of loose particles and there were no moving parts. The temperature of various parts of the cylindrical shell varied slowly, and within the range of -25 and 75°C, there was too small a temperature excursion to generate perturbations on the micrometeorite detector. The indication of the diurnal variation of the type expected strongly points to the validity of the data. The value of 10^4 tons/day as the accretion rate is similar to estimates made from zodiacal light measurements[8, 9].

5. CONCLUSIONS

Significant measurements of the impact rate of micrometeorites upon the earth were made for the first time from a satellite on Alpha 1958. An impact rate of $8{\cdot}4 \times 10^{-3}$ m^{-2} sec^{-1} was measured. From the assumptions of velocity of impact and effective momentum transfer of the impacts, the mass influx upon the earth is $1{\cdot}4 \times 10^{-11}$ g/m^2/sec for the component of cosmic dust measured. By making a further assumption for the mass distribution of interplanetary material, the daily accretion rate of the earth is found to be approximately 10,000 tons/day. The measurements also indicate the presence of a diurnal variation. Large variations from day to day have been observed. In one case, a possible micrometeorite "shower" may have occurred, although the radiant is unknown.

The total sample, though significant, is very small. Actually, this information refers to micrometeorite impacts for nearly a two-week period in the month of February. Still unknown and requiring solution by direct measurements from satellites, are a large number

of problems associated with interplanetary material at 1 astronomical unit from the sun. A good statistical sampling of micrometeoroids to determine their mass distribution, shower effects, daily and seasonal variations, is of interest. Velocities and directions may be measured to determine their orbits. Much more information is needed to understand cratering, penetration and other effects associated with impacts at velocities as high as 70 km/sec. Many more problems exist and may be answered in a direct manner using the satellite of the earth as a laboratory.

Acknowledgments—This experiment was carried out with support of the National Science Foundation and the United States Committee for the International Geo-physical Year. The Army Ballistic Missile Agency and the Jet Propulsion Laboratory prepared and launched the satellite. F. K. Dearborn, E. Baker and B. Sensinig carried through the data reduction. W. M. Alexander and, in particular, Professor J. L. Bohn contributed greatly in the design, construction and calibration of the equipment used in Alpha 1958. To those mentioned and to the many others who helped in carrying out this experiment, it is with pleasure and thanks that their cooperation is acknowledged.

REFERENCES

1. M. DUBIN, *Rocket Exploration of the Upper Atmosphere*, p. 26. Pergamon Press, London (1954).
2. M. DUBIN, *Vistas in Astronautics* (Edited by M. ALPERIN) Vol. 2. Pergamon Press, London (1958).
3. F. L. WHIPPLE, *Vistas in Astronautics* (Edited by M. ALPERIN) Vol. 1. Pergamon Press, London (1958).
4. E. MANRING and M. DUBIN, *Satellite Micrometeorite Measurements*. American Physical Society, National Academy of Science, Washington, D.C. (May 1958).
5. M. DUBIN, Tenth General Assembly, International Astronomical Union, Moscow, August 1958.
6. M. C. KELLS *et al.* *Experimental Techniques Developed for Impact Studies of Microparticles*. Third Symposium for Hypervelocity Impact, Oct. 7-9, 1958, Chicago (in publication).
7. F. L. WATSON, *Between the Planets*. Harvard University Press, Cambridge (1956).
8. S. C. VAN DE HULST, *Astrophys. J.* **105**, 471 (1947).
9. C. W. ALLEN, *Mon. Not. R. Astr. Soc.* **106**, 137 (1947).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 130–132. Printed in Great Britain.

NARROW HYDROGEN EMISSION IN THE AURORA

J. M. MALVILLE*

High Altitude Observatory, University of Colorado, Boulder, Colorado

(*Received* 11 *November* 1959)

Abstract—In the southern auroral zone during the winter of 1957, H_α was detected which had a width equal to or less than 8 Å. This observation is viewed as evidence for proton bombardment without the usual high velocity auroral particles. The alternate explanation of enhanced fluorescence of interplanetary hydrogen is discussed.

1. INTRODUCTION

The observation by Meinel[1] of Doppler shifted hydrogen lines in aurorae was the first direct evidence that high velocity protons penetrate to auroral heights in the atmosphere during times of auroral displays. Subsequent observations especially by Russian workers using higher dispersion spectrographs have confirmed Meinel's results. Galperin[2] using a spectrograph with a dispersion of 85 Å/mm has reported consistently similar H_α profiles at equal distances from the magnetic zenith. At the magnetic zenith he found the maximum of the H_α profile shifted 300–400 km/sec to the violet and a violet wing extending to 2000 km/sec. At the magnetic horizon Galperin's H_α profile had a half width of 13 Å corresponding to a velocity of 300 km/sec. All of the hydrogen profiles reported in the literature have been similarly broad and shifted. Montalbetti[3] observing with a scanning spectrometer at Churchill observed moderately bright hydrogen emission (500 rayleighs) during a period of complete absence of visible aurora. The H_α line in this enhanced airglow was less shifted to the violet than in normal aurorae indicating a maximum velocity of approach of 1200 km/sec.

The normal profile of auroral H_α in the magnetic zenith cannot be reconciled with monoenergetic protons entering the earth's atmosphere along magnetic lines of force. Hence Chamberlain[4] and Shklovsky[5] suggested a velocity dispersion of auroral protons with a distribution of the form v^{-2}. Additional evidence for low velocity protons may be inferred from Galperin's[2] observed correlation between high altitude type *A* aurorae and hydrogen emission.

Prokudina[6] has observed another type of hydrogen emission in the night sky during periods of low magnetic and auroral activity. This night sky hydrogen is significantly narrower than auroral hydrogen (half width less than the instrumental width of 2 Å). Brandt and Chamberlain[7] and Shklovsky[8] conclude that the intensities of 5 to 20 R observed by Prokudina may be explained by resonant scattering of sunlight by interplanetary hydrogen.

2. AN UNUSUAL TYPE OF HYDROGEN EMISSION

During the Antarctic winter of 1957, 507 auroral spectrograms were obtained with the patrol spectrograph at Ellsworth IGY Station (geomagnetic latitude 67° S). On all but one of the 261 spectrograms which contained hydrogen lines, H_α and H_β were typically broadened and shifted to the violet. However, during the night of August 18–19 a 430 minute exposure contained an H_α line which had a width equal to or less than the instrumental

* Fellowship Student, Institute for Solar-Terrestrial Research, High Altitude Observatory of the University of Colorado.

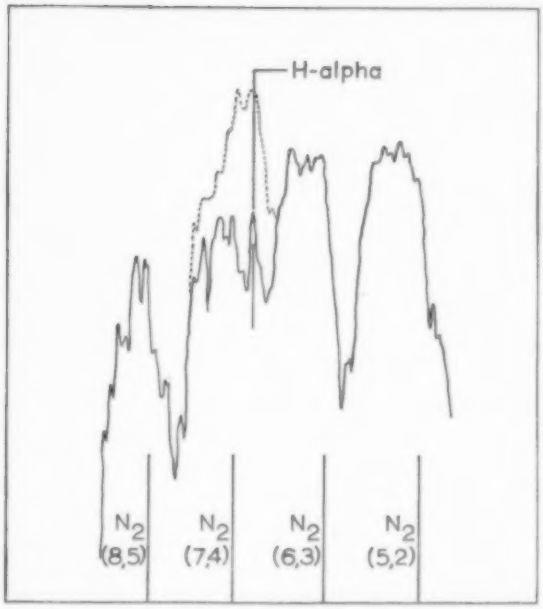


Fig. 1. Microphotometer tracing of part of spectrogram of August 18–19, 1957. The dotted line shows the approximate shape of normal broadened H_α.

width of 8 Å. Part of a microphotometer tracing of the spectrogram is shown in Fig. 1. For comparison a dotted line shows a typically broad H_α. The narrow H_α had approximately the same intensity as the Na 5890/96 doublet which in winter in polar latitudes has an emission rate of 50–70 R (Barbier and Pettit[9]).

3. DISCUSSION

Because of the low dispersion of the patrol spectrograph (300 Å/mm), it is difficult to exclude the possibility that the narrow H_α line came from a high altitude arc viewed on the magnetic horizon. However, the uniqueness of the line is very suggestive of a significant departure from normal auroral conditions. On the 260 other spectrograms, even on the magnetic horizon, H_α was more broad (15 Å half width). When the first positive band system of N_2 was present, the violet wing of normal H_α was always blended with the 7,4 band head. Emission in the narrow H_α line extended from the southern horizon to 45° above the northern horizon. As the sky was overcast during a portion of the exposure, the radiation cannot be identified definitely as having originated at the magnetic zenith or the magnetic horizon. However, even the low velocity protons observed by Montalbetti[3], which must have been stopped in the atmosphere above 130 km, yielded H_α on the magnetic horizon with a half width of 15 Å. Consequently the protons responsible for the 8 Å wide H_α line must have had a velocity considerably less than 1200 km/sec.

H_α with a width of 8 Å or less at the magnetic zenith may be explained by a flux of protons with a velocity less than 200 km/sec and an energy less than 0·2 keV. According to recent calculations by Chamberlain* a 500 km/sec proton would radiate 0·18 quanta of H_α in its passage through the atmosphere before it is brought to rest. If we assume a mono-energetic beam of protons, an emission rate of 70 R corresponds to a proton flux of approximately 4×10^8 protons/cm² sec. Velocities of 500 km/sec and densities of 1–10^3 particles/cm³ are suggested for solar streams originating from *M*-regions. The observed narrow H_α may have been produced by such low velocity corpuscular streaming which apparently represents a normal condition of the declining sun. Such narrow hydrogen emission may be common across the polar cap but would be easily missed because of its faintness and the impossibility of detecting it when normal broad hydrogen emission is present. This observation appears to complement the conclusions of Chamberlain[4], Galperin[2], and Shklovsky[5] that low velocity protons are associated with the 2000 km/sec protons observed in normal aurorae.

The existence of such low energy protons in the auroral zone again brings up the problem which has been recognized in geomagnetic phenomena for some time, namely, the existence of particles of energies far below the geomagnetic cut-off energy. Particles such as those discussed here and those associated with

* Private communication.

type III polar cap absorption (Leinbach and Reid[10]) suggest that at times of even moderate magnetic disturbances solar particles experience little or no transverse magnetic field while penetrating into the polar cap.

Because Ellsworth is in the southern auroral zone, it seems natural to explain the narrow H_α line by a beam of low energy protons. An alternate explanation would be an enhanced resonant scattering of sunlight by interplanetary hydrogen. The total number of interplanetary/hydrogen atoms responsible for scattering an intensity I_α of H_α is

$$N_H = \frac{4\pi I_{32}}{g_{32}} = \frac{4\pi I_{32}}{\left(\frac{\pi e^2}{mc}\right) f_{13} \pi F_{13} \left\{\frac{A_{32}}{A_{32}+A_{31}}\right\}} \quad (1)$$

where f_{13} is the oscillator strength for Lyman-beta, the A's are the Einstein coefficients, and F_{13} is the flux of solar Lyman-beta. We adopt Brandt and Chamberlain's[7] value for g_{32} which was obtained assuming no absorption core of the solar line. With $g_{32}=2{\cdot}3\times10^{-6}$ photon/atom sec and $4\pi I_{32}=7\times10^{7}$/photons cm^{-2} sec^{-1}, the number of hydrogen atoms along the line of sight is $N_H=3\times10^{13}$ atoms/cm^2. This value is a factor of ten greater than that indicated by the observations of Kupperian *et al.*[11] on Lyman-alpha in the night sky. We can not attribute an apparent ten fold increase in N_H to an increase in the flux of solar Lyman-alpha because the sun was relatively quiet during this period. Instead, such an increase may result from a density inhomogeneity of the interplanetary medium. It is unlikely that an enhancement with a life time of two or three days in the neutral hydrogen density surrounding the earth would be produced by low velocity corpuscular streaming from the sun. The recombination time for electrons and protons after ejection from the sun is approximately

$$\tau \simeq (N_p \alpha_{rec})^{-1} \quad (2)$$

with the proton density in the stream $N_p=10^4$/cm^3 and the recombination coefficient, $\alpha_{rec}=10^{-11}$, the recombination time is approximately 100 days.

4. SUMMARY

The observation of H_α with a width of 8 Å or less may be explained by (1) a stream of low energy protons incident upon the atmosphere or (2) scattering by interplanetary hydrogen. Alternative (2) involves no serious conceptual difficulties; but because the observation was obtained in the auroral zone, explanation (1) appears the more reasonable. So interpreted, the observation represents direct evidence for the existence of low velocity auroral protons, already inferred from the normal H_α profile and from the association of H_α with high altitude type *A* aurorae.

Acknowledgment—The author wishes to express his thanks to D. E. Billings, J. W. Chamberlain and J. W. Warwick for reading the manuscript.

REFERENCES

1. A. B. Meinel, *Astrophys. J.* **113**, 50 (1951).
2. Iu. I. Galperin, *Soviet Astronomy*, **1**, 133 (1957); *Planet. Space Sci.*, **1**, 57 (1959).
3. R. Montalbetti, *J. Atmos. Terr. Phys.* **14**, 200 (1959).
4. J. W. Chamberlain, *Astrophys. J.* **126**, 245 (1957).
5. I. S. Shklovsky, *Bull. Akad. Nauk SSSR*, **81**, 367 (1951).
6. V. S. Prokudina, *Spectral, Electrophotometrical and Radar Researches of Aurorae and Airglow*, p. 43. Academy of Sciences of the U.S.S.R., Moscow (1959).
7. J. C. Brandt and J. W. Chamberlain, *Astrophys. J.* **130**, 670 (1959).
8. I. S. Shklovsky, *Planet. Space Sci.* **1**, 63 (1959).
9. D. Barbier and H. Pettit, *Ann. de. geophys.* **8**, 232 (1953).
10. H. Leinbach and G. Reid, *Phys. Rev. Letters*, **2**, 61 (1959).
11. J. E. Kupperian, Jr., E. T. Byram, T. A. Chubb and H. Friedman, *Planet. Space Sci.* **1**, 3 (1959).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 133–149. Printed in Great Britain.

FURTHER STUDIES OF "SPREAD-*F*" AT BRISBANE–I

EXPERIMENTAL

G. G. BOWMAN

Physics Department, University of Queensland, Brisbane

(*Received* 10 *September* 1959)

Abstract—Rotating-spaced loops at Brisbane recorded the azimuths-of-arrival of pulsed 3·84 Mc/s signals from Armidale (202°, 355 km from Brisbane); also pulsed signals at the same frequency, from a transmitter on the site of the loops.

These and other normal-incidence records indicate that the irregularities of the *F2*-layer responsible for "Spread-*F*" records at Brisbane, are ripples of considerable lateral extent with a "wavelength" varying from 20 km to over 100 km. The position of the ripple in each ionization contour changes with height in such a manner as to suggest a simple relationship with the geomagnetic field orientation.

Seasonal variations of the phenomenon show a winter maximum and a summer sub-maximum, and there is a distinct inverse sunspot-cycle relationship. The type of spreading defined as range-spreading generally has a peak of occurrence around midnight, and that defined as frequency-spreading peaks during the sunset and sunrise periods, after an allowance is made for a diurnal variation associated with the changing f_0F2 value.

1. INTRODUCTION

The ionospheric phenomenon of "Spread-*F*" makes itself manifest on ionosonde records by additional traces, which are generally similar in shape to the main *h'f* trace, but with critical frequencies (f_0F2) which may be greater than, or less than, that of the main trace. It is seen at Brisbane associated with the one-hop trace obtained at night, the frequency of occurrence depending on season and state of the sunspot cycle (McNicol *et al.*[1]; Singleton[2]).

The "Spread-*F*" records at Brisbane display a variety of forms. Observations of the satellite traces at the low frequency end of ionograms, where retardation can be neglected, indicate, at certain times, ranges more than 10 km greater than the true range of the main echo. At other periods the true range of the satellite echo is little greater than that of the main trace, but the critical frequencies are different. The first type of spreading has been called "range-spreading", and the other "frequency-spreading".

Other ionograms showing "Spread-*F*" merely exhibit diffuse traces. Evidence presented in this paper indicates that nearly always these result from satellite traces so close in true range or critical frequency that the resolution of the equipment is insufficient to separate them. It is convenient to classify spreading into the groups "resolved" and "unresolved", as illustrated in Fig. 1, on the basis of the ionograms. It must, however, be remembered that such distinction is a function of the instruments used for observation. Many borderline cases occur.

A study of resolved range-spreading, using mainly normal-incidence reflection, has already been published[1]. Further information on this phenomenon, together with unresolved range-spreading and unresolved frequency-spreading, has now been obtained employing a rotating spaced-loop direction-finding system, based on that of Ross[3]. This was located at Brisbane (lat. 27·5°S., long. 152·9°E., geomag. lat. 35·1°S., magnetic dip 57°) and observations made on reflected pulses originating from (i) a transmitter at Armidale, distant 355 km and

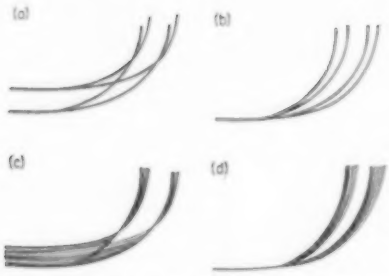


Fig. 1. Ionogram sketches of (a) resolved range-spreading; (b) resolved frequency-spreading; (c) unresolved range-spreading; (d) unresolved frequency-spreading.

bearing 202° (from true north) from Brisbane and (ii) a transmitter located close to the loops, and operating on the same frequency (3·84 Mc/s). The spaced-loop direction-finding system while it does not provide information on the elevation angle, as does that of Thomas and McNicol[4], has the advantage of allowing the azimuths-of-arrival to be determined, even when satellite echoes are closely spaced in range.

In addition, further analysis of *h'f* and other records described in Ref. (1) has been undertaken.

2. EQUIPMENT

(a) *Transmitters*

The transmitters at Armidale and Brisbane operated with a peak output of 2 kW. The pulse duration was 70 μsec and half-wave dipole aerials, 10 m above the ground were used. The Brisbane transmitter had a pulse-repetition-frequency of 50 c/s, the modulation being controlled by the mains, while the pulse-repetition-frequency of the Armidale transmitter was 333·3 c/s. This frequency was maintained to a high degree of accuracy by a crystal-controlled unit.

(b) *Receiver and displays*

A common receiver (Loran *AN/APN*-4, modified) was used for both sets of signals, the output being connected to two separate C.R.O. display units. The ground pulse from the Brisbane transmitter was used to trigger the time-base of one display, thus locking the positions of the echoes from the local transmitter. Signals from positions with virtual ranges from 0 to 600 km were recorded, and height marks at 50 km intervals were placed on the record once every 4 min.

The second display was triggered by a crystal-controlled unit, identical with the one modulating the Armidale transmitter. Small differences in the two crystal-controlled frequencies made the pattern of signals from Armidale drift slowly across the C.R.O. face. However, with a fine adjustment on the Brisbane unit, it was possible to arrest this drift so that, at times, as long as half an hour was required for the pattern to drift right across. The time-base velocity was so adjusted that, as the signal pattern disappeared on one side of the display, it reappeared on the other. Thus all signals were recorded at all times. The distance between patterns represented 900 km range and allowed a calculation of range differences between the main *F2*-trace and other traces.

35 mm film moved at right angles to the time base, on which the signals were represented by a blackening out of the trace, at a rate of 3·8 cm/min and a small neon light indicated on the film a particular orientation of the system once every rotation.

(c) *Direction-finder*

The rotating-loop system consisted of two loops 2 m square spaced 10 m apart, with their planes parallel to each other and perpendicular to the booms joining them (Fig. 2). A cabin (centrally situated with respect to the loops) housed the recording equipment, and a synchronous electric motor which rotated the system at a rate of one revolution per 2 min. Power and subsidiary cables were buried 3 ft underground, and entered the cabin through slip-ring contacts.

The loop aerials, consisting of two turns of coaxial cable, were tuned to resonance, and balanced for phase and amplitude by adjustments at the base of each loop. The aerial outputs were brought by coaxial cables into

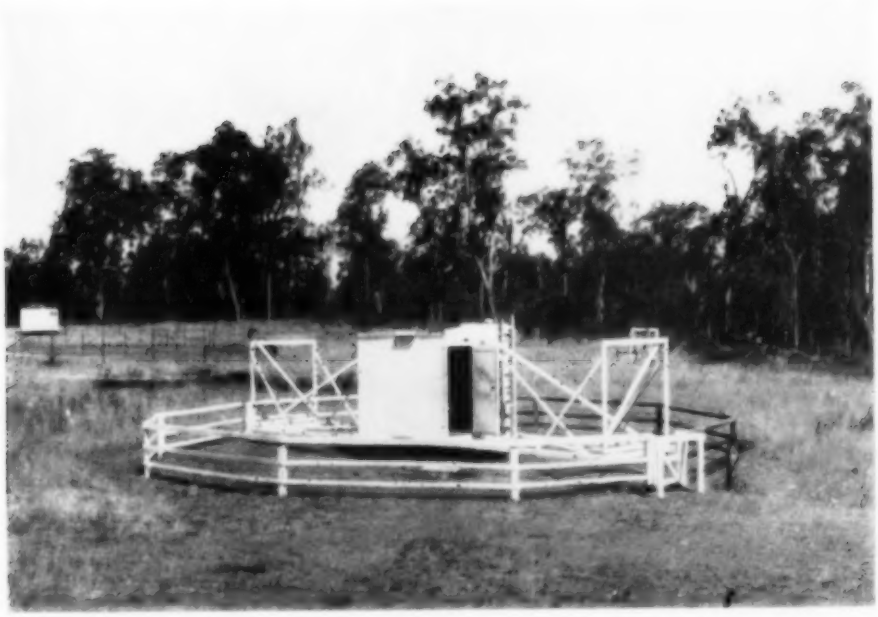


Fig. 2. Rotating spaced-loop direction-finding system.

Facing page 154

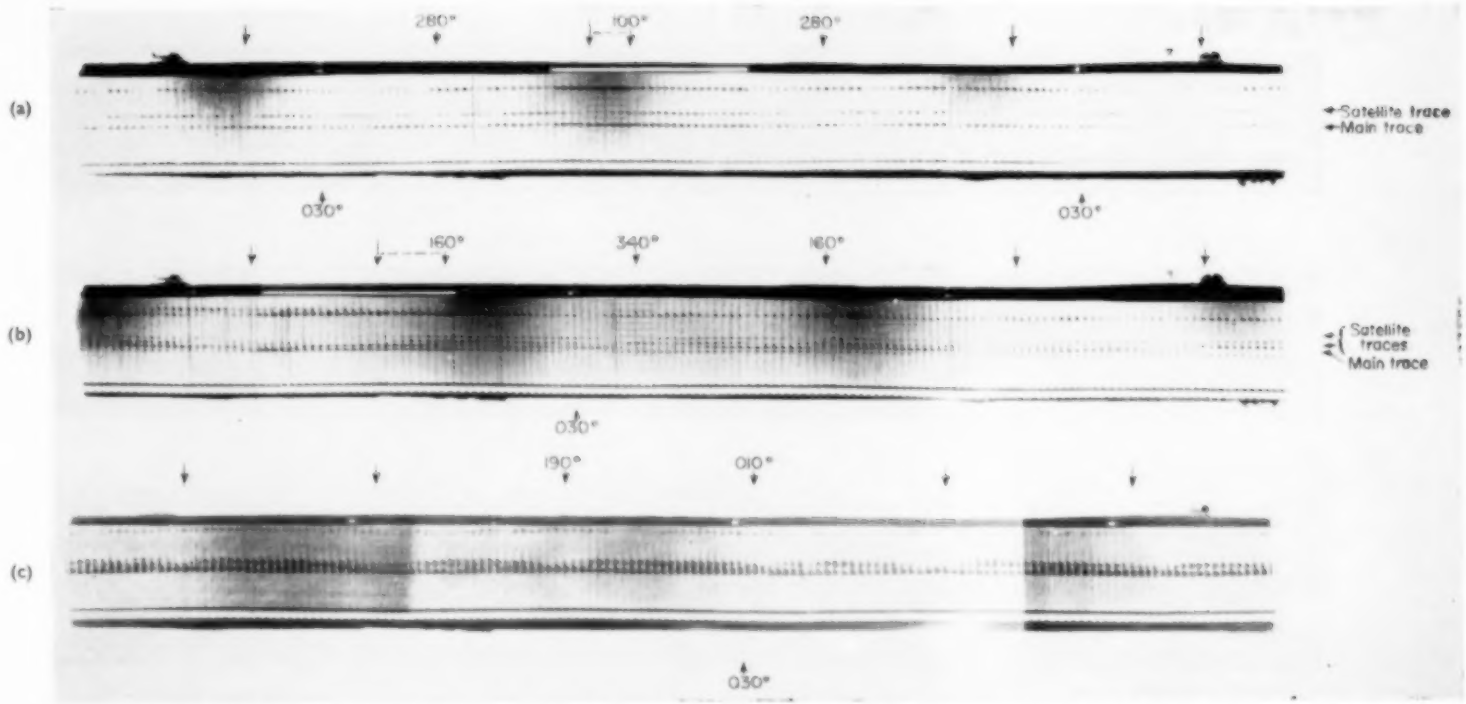


Fig. 4. Normal-incidence rotating-loop records indicating azimuths-of-arrival. (a) Resolved range- spreading; single satellite trace at 21.40 on 7 August 1958; (b) Resolved range-spreading; multiplicity of satellite traces at 21.50 on 9 August 1958; (c) Unresolved range-spreading at 22.10 on 25 July 1958.

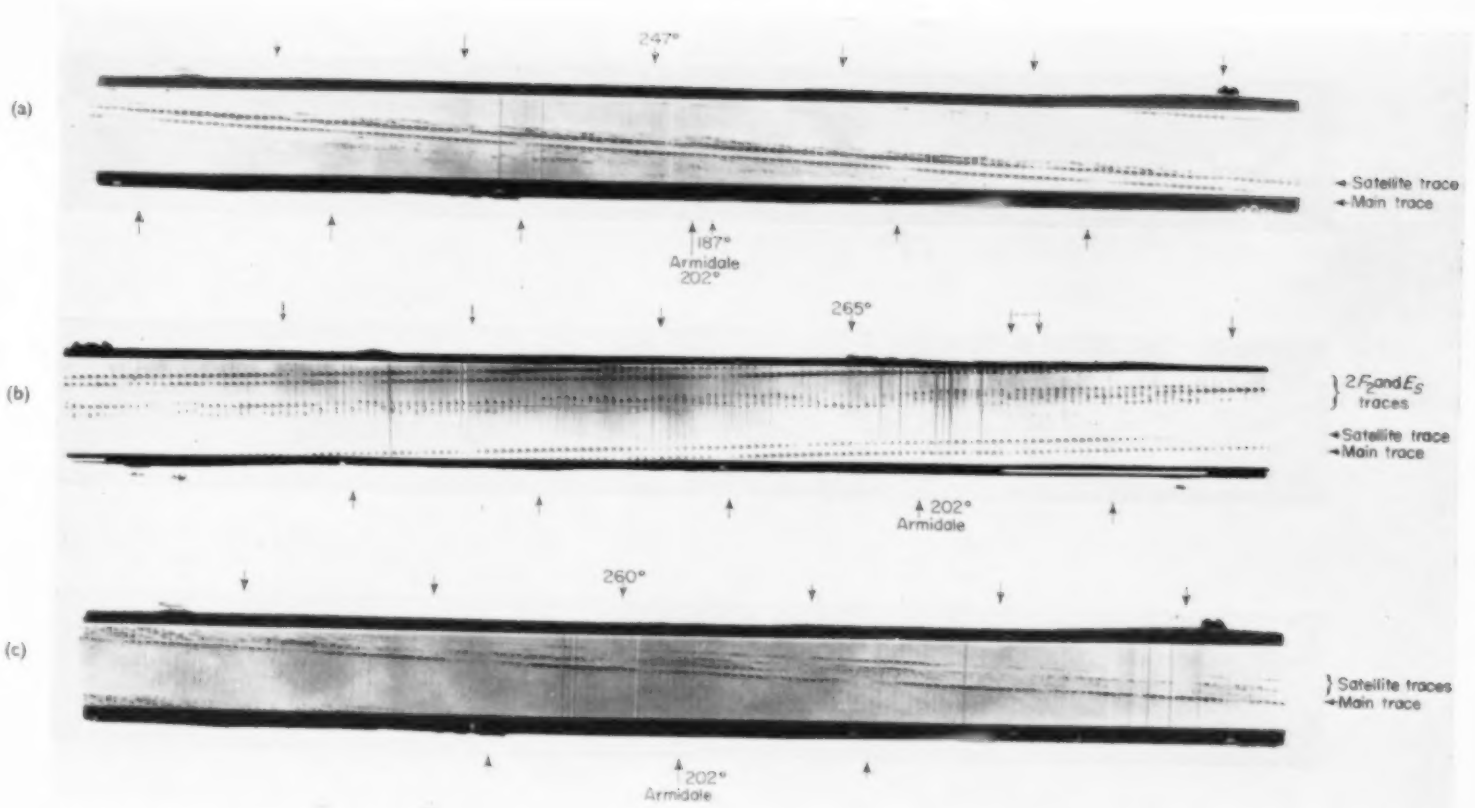


Fig. 5. Oblique-incidence rotating-loop records indicating azimuths-of-arrival. (a) Resolved range-spreading at 20.20 on 8 July 1958; swept-gain range 20 dB.; (b) Resolved range-spreading at 22.30 on 4 August 1958; additional nulls due to fading; swept-gain range 45 dB.; (c) Resolved range-spreading at 21.00 on 8 July 1958; multi-plicity of satellite traces; swept-gain range 20 dB.

VOL
2
1959

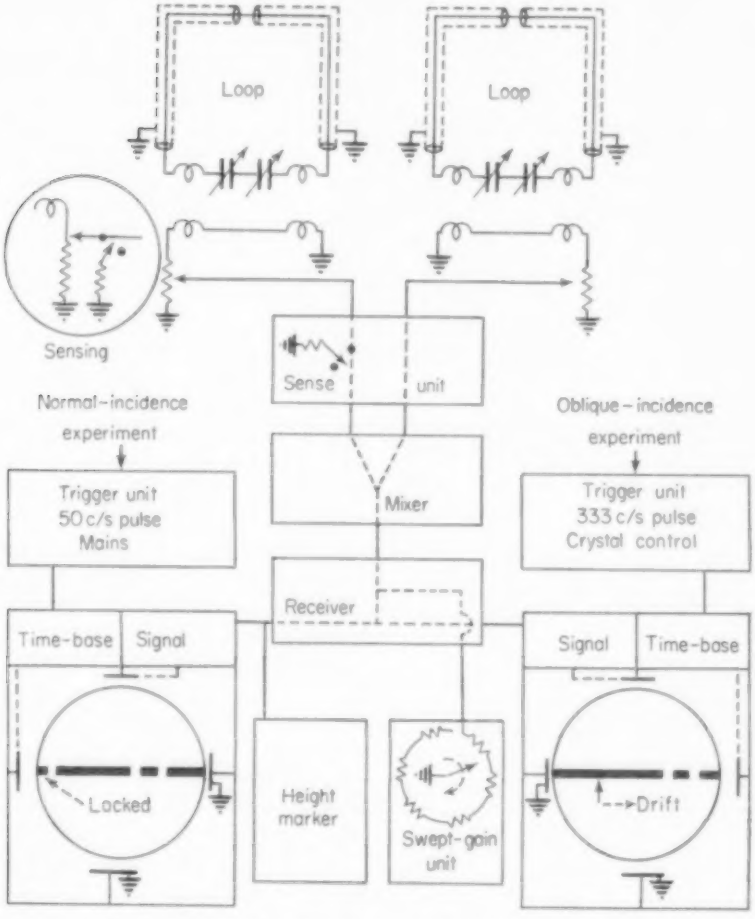


Fig. 3. Schematic diagram of rotating-loop system equipment.

the cabin where their signal strengths were compared in a difference unit (Fig. 3). Its output gave a "null" response when the boom orientation was at right angles to the azimuth-of-arrival of the signal.

The ambiguity of direction was resolved by switching in a suitable resistance between one aerial output and earth so that the amplitude balance of the aerials was disturbed. This had the effect of rotating the position of each "null" by about 30°, the displacement of the "nulls" being in the opposite sense. The resistance was switched in for one min in six. The neon light was also used to indicate on the film when the "nulls" were displaced.

Since the "nulls" provided by the spaced-loop system, with the signals coming in at high elevations, were far from perfect it was arranged for the gain of the receiver to be varied over a range 48 times every rotation, i.e., once every 2·5 sec. Originally a method similar to that described in Ref. (1) with a 20 dB variation of sawtooth type was employed. Various other arrangements were used until 25 July 1958, when the gain variation was made 45 dB in 9 equal steps of 5 dB.

Figure 4(a) gives an example of a normal-incidence satellite trace showing the appearance of the "nulls", and the displacement of one "null" to resolve the ambiguity of azimuth. "Nulls" displaced to the right indicate the true direction-of-arrival when the right-hand edge of the neon direction indicator is called 030° (true). Satellite traces on oblique-incidence records are shown in Fig. 5 (a) and (b). Fig. 5(b) is complicated by additional "nulls" due to fading. However, the loop "nulls" can be distinguished by their repeated occurrence.

(d) *General*

A block diagram of the recording equipment arrangement is shown in Fig. 3. The oblique-incidence experiment gave suitable records from 20 June 1958 to 17 August 1958 (inclusive). Recordings were made mainly during the period 20.00–24.00 hr. However, occasional records were made for the rest of the night and during the sunset and sunrise periods.

3. EXPERIMENTAL RESULTS

(a) *Plan position of irregularities*

(i) *Resolved range-spreading.* Well-defined satellite traces were recorded by both the oblique-incidence and normal-incidence experiments. During the 120 hr of concurrent recording (20 June 1958 to 17 August 1958 inclusive) satellite traces appeared in the record of signals from Armidale on thirty-two occasions when discrete satellite traces appeared on the normal-incidence recordings, and five occasions when they were absent on the normal-incidence recordings. Also, during this interval, twenty-five normal-incidence satellite traces occurred when no satellite trace appeared in the Armidale records. These satellites were nearly all divergent types on both sets of records (cf. Ref. 1 for classification of satellite types). A total of four convergent type and one mixed type was found at normal-incidence and one convergent type and one mixed type at oblique-incidence. This is what would be expected from the previous study (Fig. 23 of Ref. 1) as the satellite traces are predominantly divergent in the hours before midnight, when most of the rotating-loop records were taken.

Figures 4 and 5 show sample records. In Fig. 5(a), what would normally be called the main trace is displaced 15° towards the east, while the satellite trace shows an extra range of 125 km, and a displacement of 45° to the west. The displacement of 15° towards the east was the highest ever observed for the first *F2*-trace received. Displacements over 7½° were exceptional.

Assuming retardations of the satellite signals (oblique- and normal-incidence) to be the same as those of the main trace (oblique- and normal-incidence), true ranges were calculated using data from ionograms. It was also assumed that the height of reflection for the satellite trace was the same as for the main trace. These data, together with the direction-of-arrival information, allowed a determination of the position and orientation of the reflecting surface, if it is further assumed that reflections were specular and no refraction took place.

Two examples of the plotting method used and the results obtained are shown in Fig. 6(a) and (b). In Fig. 6(a) the additional echoes are satellites of one-hop *F2*-reflections, whereas in Fig. 6(b) the satellites are related to the two-hop *F2*-reflections. Fig. 6(b) illustrates the large lateral deviation sometimes encountered.

Despite the many assumptions made here, the correlation between the orientation of oblique-incidence and normal-incidence reflecting surfaces was sufficiently good to establish that the irregularities have a lateral extent of at least several hundred kilometres.

Rotating-loop records for the night of 10–11 July 1958 show satellite echoes occurring for several hours for both normal-incidence and oblique-incidence equipments. One divergent satellite on the normal-incidence record was suitable for a calculation of the approximate speed of movement perpendicular to the frontal irregularity. This was deduced from the rate of change of range (cf. Ref. 1), and gave a speed of 90 km/hr in the direction 310°. For the plot which followed, the ionospheric irregularities were imagined to remain stationary while

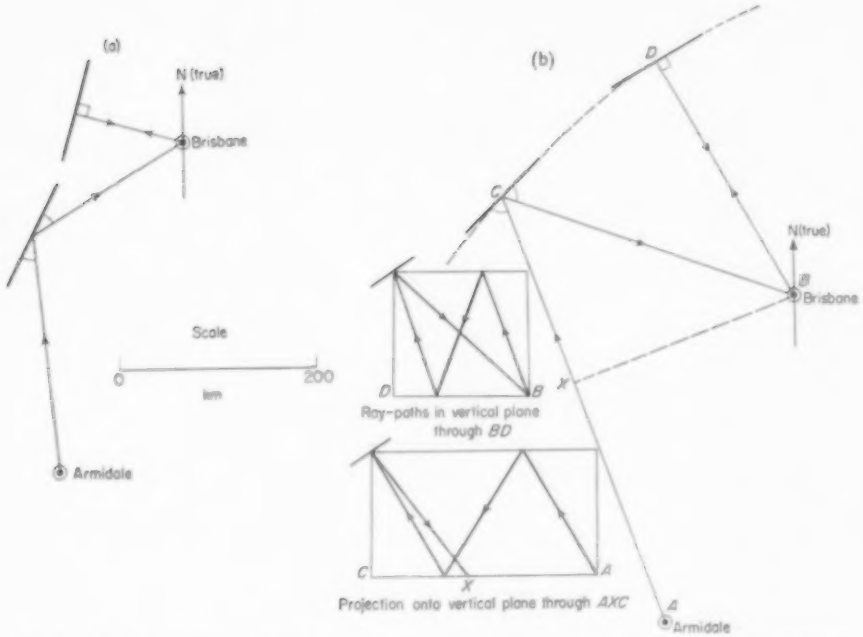


Fig. 6. (a) Plan position of reflecting surfaces causing resolved range-spreading at 21.20 on 7 August 1958; (b) Plan position of reflecting surfaces causing resolved range-spreading associated with the two-hop F_2 trace at 19.52 on 11 July 1958.

the plotting positions of Armidale and Brisbane moved with the velocity of the configuration. The previous calculation gave the component of velocity in the direction 310° but the true velocity was not known. Velocities were tried in several directions taking into account the component in the direction 310°. Positions and orientations of reflection surfaces were then plotted in the manner of Fig. 6 from the appropriate position of each station as it moves relative to the frontal configuration. For most directions the distribution of reflection surfaces was random, but as the direction approached that of 010° some order came into their positions and orientations. Fig. 7 shows a plan of the surfaces when a velocity of 010°, 175 km/hr is used. It indicates a series of fronts, of considerable lateral extent, with a spacing between them of the order of 100 km.

An azimuth reading on a normal-incidence record indicates the vertical plane perpendicular to the front. Fig. 8(a) shows the distribution of these perpendiculars (twenty cases) for resolved range-spreading as from 3 August 1958, when the sensing device was installed. Virtually all are within the semi-circle from 240° through 360° to 060°. Earlier results have a 180° ambiguity. Fig. 8(b) shows a plot of all readings (seventy-nine cases) assuming none to lie between 060° and 240°. 95 per cent of satellites used in this distribution are of the divergent type so that nearly all indicate the apparent direction of travel of the fronts. Fig. 19 of Ref. (1) is a plot of component velocities perpendicular to the fronts and here it is found that most of the points lie in the NW quadrant. Fig. 8, therefore, appears to be consistent with this previous investigation.

(ii) *Unresolved range-spreading.* This type of spreading was classified as unresolved because of its appearance on ionograms. It

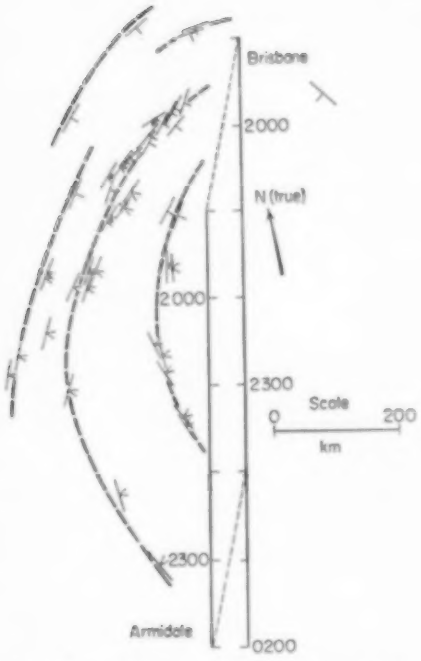


Fig. 7. Plan position of reflecting surfaces of ionospheric disturbance on night of 10–11 July 1958.

was pointed out[1] that a swept-gain record gives better resolution than a fixed-gain record, other things being equal. The rotating-loop normal-incidence record, with its fast swept-gain (period 2·5 sec), showed resolution of nearly all "Spread-*F*" echoes, even when they appeared as unresolved on the routine ionograms. Fig. 9(a(i)) shows an example of this type of recording, compared with a swept-gain record at 2·28 Mc/s, Fig. 9(a(ii)), taken at the same time but having a swept-gain period of 1 min. The relevant ionogram is also shown. Another important feature of this type of spreading, as recorded by the rotating-loop system, is the appearance of nulls on practically every occasion. Fig. 4(c) illustrates these nulls. Sometimes as many as ten satellite traces can be seen at the one time. These records suggest frontal irregularities similar to that illustrated in Fig. 7, but having a closer spacing.

Estimates have been made of the frontal spacings using the range separation of the satellite traces for the period 25 July 1958 to 26 August 1958 inclusive. The distribution is shown in Fig. 10, and reveals that, at least for the period investigated, the most frequently occurring frontal separation is 35 km. Fig. 11(b) shows a structure with frontal separation of 25 km, using information derived from a normal-incidence record.

A similar multiplicity of fronts is sometimes exhibited in resolved range-spreading. Thus Fig. 4(b) illustrates a configuration with three satellite traces with frontal spacings of approximately 80 km. Fig. 5(c) shows an example of oblique-incidence rotating-loop recording where several satellite traces are recorded at the same time. Fig. 11(a) is a plot using this record and the normal-incidence record showing a frontal separation of 40 km.

The distribution of vertical plane perpen-

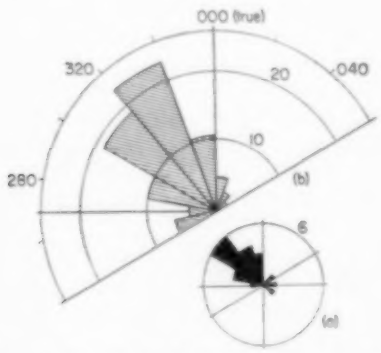


Fig. 8. Azimuths-of-arrival of signals for resolved range spreading; (a) Sensing in operation; (b) 180° ambiguity in direction.

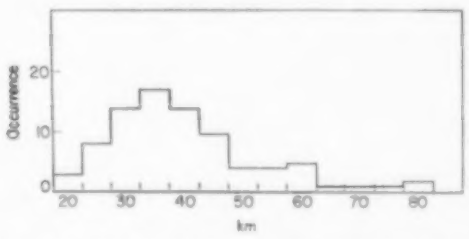


Fig. 10. Distribution of frontal separations for unresolved range-spreading irregularities for th period 25 July 1958 to 26 August 1958.

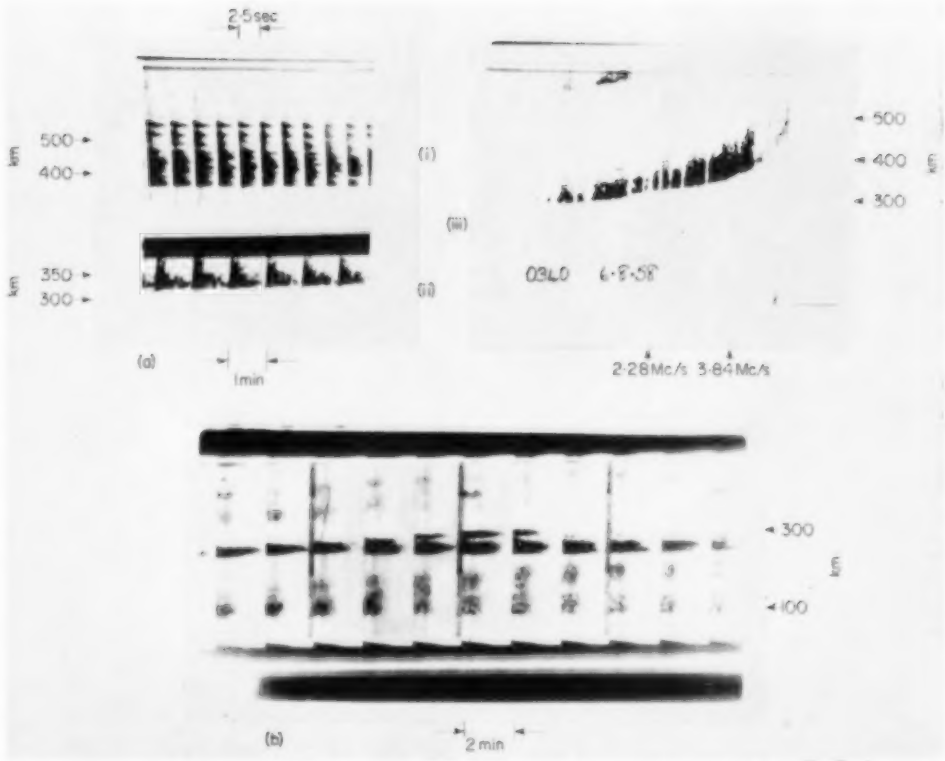


Fig. 9. (a) Unresolved range-spreading at 03.4 on 6 August 1958 ; (i) Rotating-loop *h't* record at 3·84 Mc s ; swept-gain period 2·5 sec ; (ii) Normal *h't* record at 2·28 Mc s ; swept-gainperiod 1 min ; (iii) Ionogram of spreading. (b) Swept-gain h't record (period 2 min) at 20.15 on 29 June 1955 showing enhancement of satellite signal.

Facing page 138

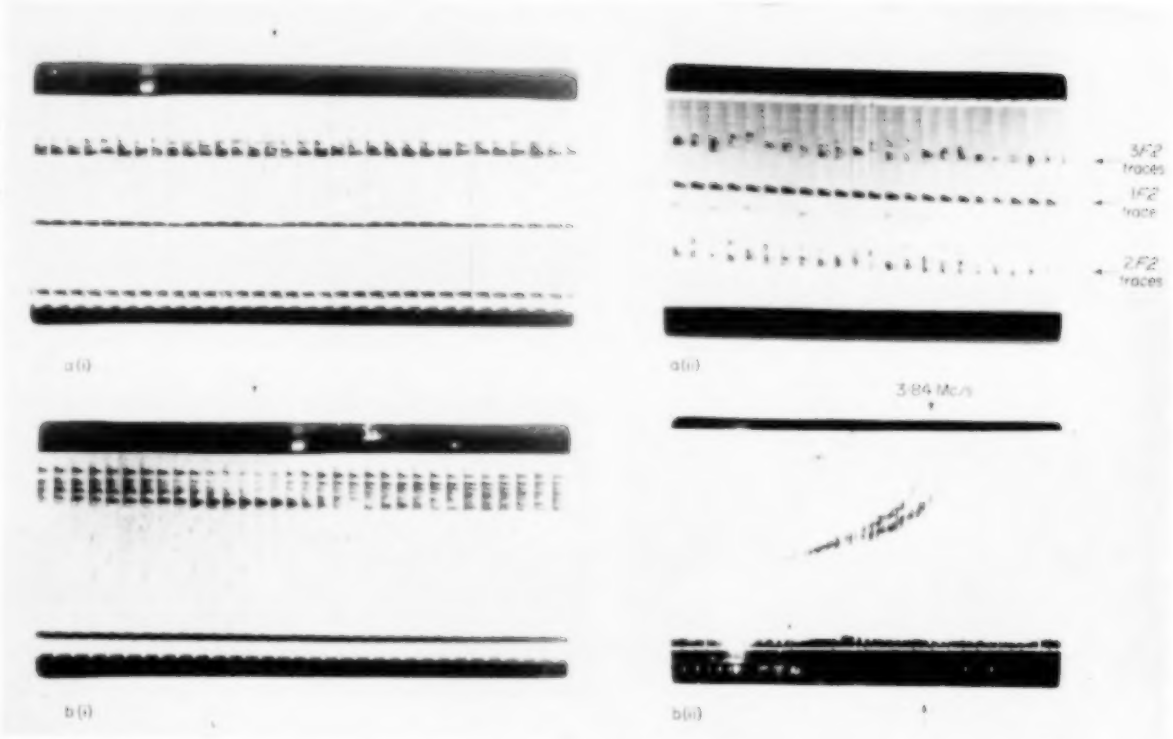
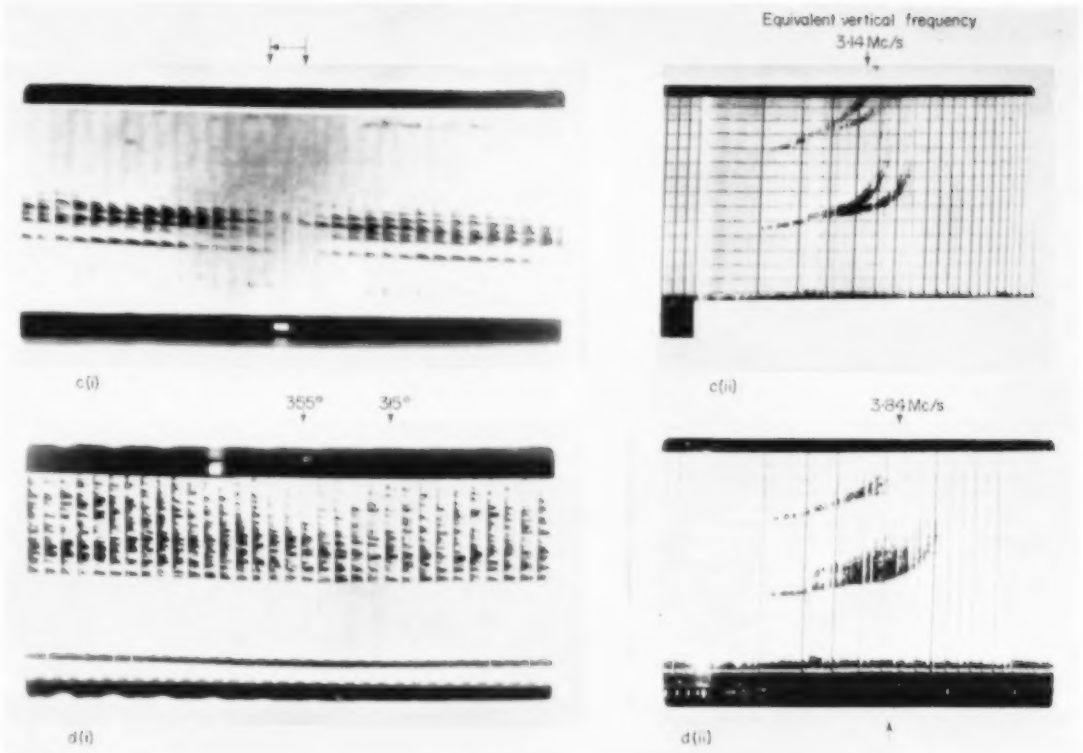


Fig. 14. (a) Two-hop unresolved range-spreading ; (i) normal-incidence rotating-loop record at 20.30 on 23 August 1958 ; (ii) oblique-incidence rotating-loop record at 20.20 on 15 August 1958.
(b) Unresolved frequency-spreading at 23.50 on 1 August 1958 ; (i) normal-incidence rotating-loop record ; (ii) ionogram.



(c) Unresolved frequency-spreading at 05.30 on 18 July 1958 ; (i) oblique incidence rotating-loop record ; (ii) ionogram.
(d) Complex unresolved spreading at 05.30 on 30 July 1958 ; (i) normal-incidence rotating-loop record ; (ii) ionogram.

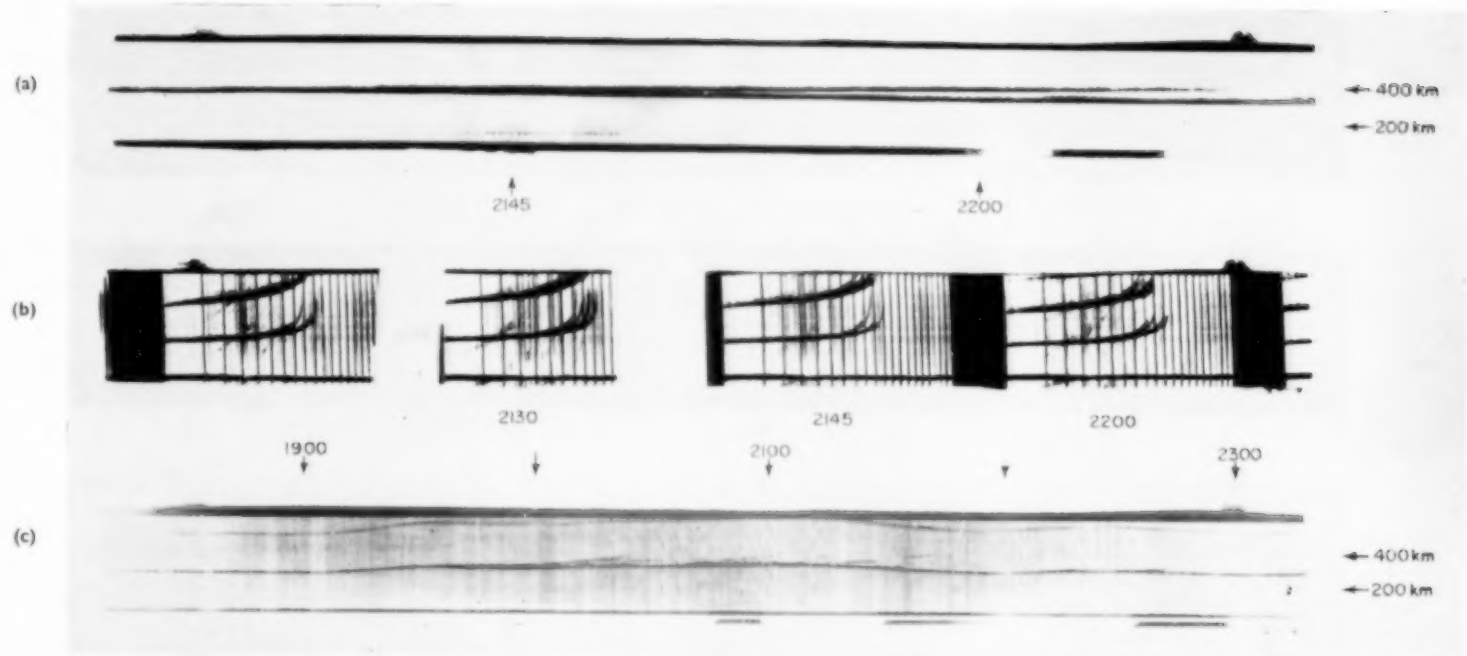


Fig. 15. Normal-incidence "Spread-*F*" records at Brisbane. (a) Resolved range-spreading satellite trace; fixed frequency (3·84 Mc s) *h't* record on 30 June 1958; (b) Resolved frequency-spreading; *h'f* ionograms on 17 December 1944; (c) Resolved range-spreading; fixed frequency (3·84 Mc s) *h't* record on 7 August 1958.

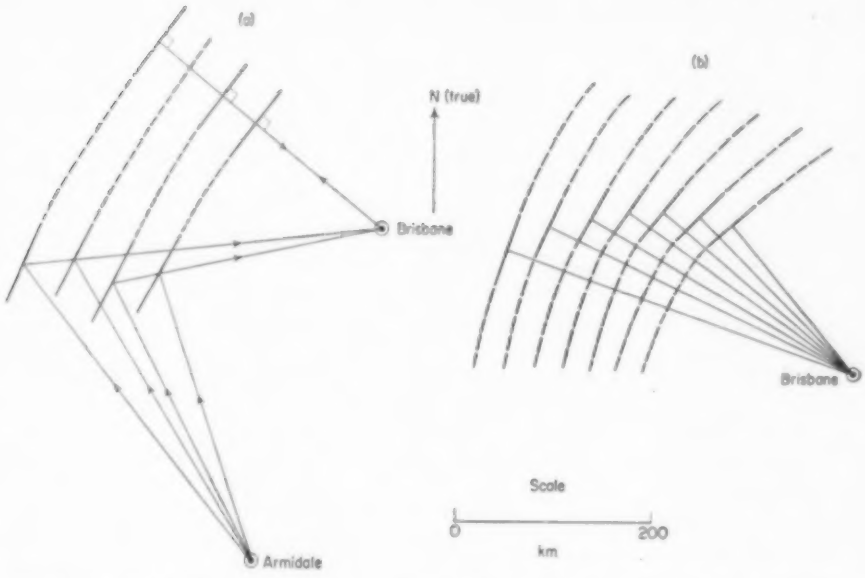


Fig. 11. (a) Plan position of reflecting surfaces at 21.00 on 8 July 1958; (b) Plan position of reflecting surfaces causing unresolved range-spreading at 20.00 on 27 July 1958.

diculars to the fronts for unresolved range-spreading is shown in Fig. 12. Comparison with the distribution found for resolved range-spreading shows a shift towards the north.

(iii) *Resolved frequency-spreading.* Using direction-of-arrival records at 2·28 Mc/s[4] it was possible to find elevations and bearings for certain satellite traces classified as resolved frequency-spreading. This was possible only when the group retardation at this frequency separated the satellite trace adequately from the main trace. The average elevation for the ten cases found was 87°. The azimuths are shown plotted in Fig. 13.

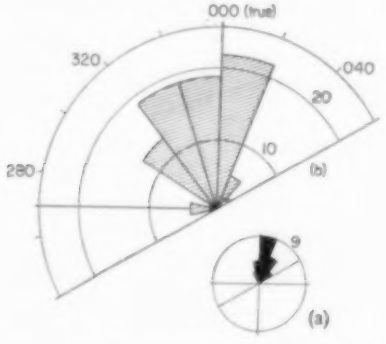


Fig. 12. Azimuths-of-arrival of signals for unresolved range-spreading; (a) Sensing in operation; (b) 180° ambiguity in direction.

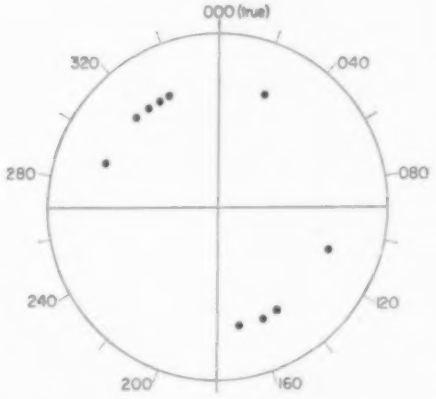


Fig. 13. Azimuths-of-arrival of signals for resolved frequency-spreading.

(iv) *Unresolved frequency-spreading.* Occasionally, especially in the early morning, the f_0F2 value fell to a sufficiently low value to permit unresolved frequency-spreading (identified by reference to ionograms), to be observed on the rotating-loop records, both normal-incidence and oblique-incidence. Here again, there is evidence of resolved traces and, on the normal-incidence records, nulls indicate that the signals come from a definite direction (Fig. 14b). Fig. 14(c) shows an oblique-incidence record where the azimuth-of-arrival of the satellite *O*-ray echo swings from the direction of Armidale to about 20° west of this as between the five satellite traces present. Such records suggest irregularities very similar to those found for range-spreading.

Fig. 14(d) illustrates a record showing complex spreading. The rotating-loop record can be divided into three parts. With increasing range there is first the main *x*-ray with very little frequency-spreading; secondly, the main *O*-ray with considerable unresolved frequency-spreading from the direction 315°; and thirdly, from a greater range, and showing pronounced directional effect, range-spreading echoes from 355°.

(b) *Vertical section of irregularities*

(i) *Resolved range-spreading.* Fig. 15(a) illustrates a good example of a satellite trace on a *h't* record, the irregularity being overhead at 21.30 hr on 30 June 1958. Fig. 15(c) shows (on a different scale) a series of divergent satellite traces, observed over a period of about 2 hr. The speed of movement perpendicular to the irregularity front, indicated on Fig. 15(a), was calculated by a method described in Ref. (1). (N, h) profiles were determined at 10-min intervals from ionograms[5] for a period before and after the zenithal passage of the irregularity. If the frontal orientation and the direction of movement is as indicated in Fig. 16(a), these profiles will represent (N, h) curves at points *A, B, C, D, E* assuming that the reflected ray does not depart from the vertical at these points. Assuming the irregularity to have a frontal character it follows that points *F, G, H, I* have the same (N, h) profiles as points *B, C, D, E*, respectively. Fig. 16(b) shows vertical sections, through *AI*, of the surfaces of equal ionization density, compiled from (N, h) profiles, and the measured speed perpendicular to the front. The irregularity appears as a ripple in each ionization contour, the ripple being progressively displaced horizontally with increasing height. The straight lines—"slope-lines"—joining corresponding points on successive contours make an angle of approximately 25° with the vertical. We shall term this angle the slope of the irregularity. The diagram in Fig. 16(c) illustrates that, if the geomagnetic field direction is projected on to a vertical plane, perpendicular to the frontal orientation, the slope of this line is approximately the same as the slope of the irregularity. In Fig. 16(b) it is interesting to note that a small patch of E_s was found in the region defined by producing the slope-lines down to the 100 km level (assuming the whole ionosphere moved as a body).

The contour shapes developed in this way must only be approximations as no allowance has been made for reflection away from the normal and distortion of the normal-incidence ray-paths caused by the kinking of the contours. Nevertheless the position of the trough and crest in each ionization contour can be considered a good approximation as, at these positions, the contours are close to the horizontal. A further complication is the possibility that, if different frequencies are reflected with different elevations, the *h'f* curve becomes unsuitable for analysis by the Schmerling method.[5] If the contours of Fig. 16(b) are assumed to be correct ones and at a particular time perpendiculars are drawn on to each ionization contour the resulting virtual ranges will not be significantly different from those arising from the points plotted vertically at this time. It has therefore been assumed that, for relatively shallow ripples of the type illustrated by Fig. 16(b), the method gives a reasonable estimation of the shape of the cross-section of the irregularity.

One of the few converging satellites was

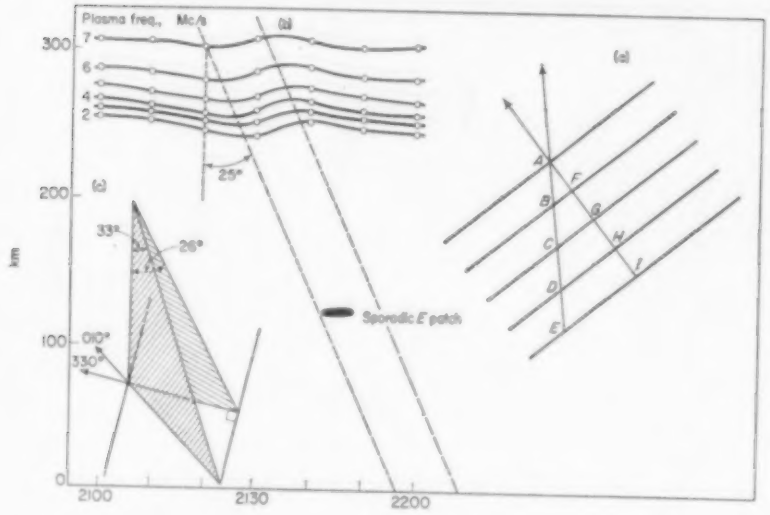


Fig. 16. (a) Illustrated movement of frontal irregularity relative to recording station at A. (b) Variation of contours of equal ionization at the irregularity causing a divergent satellite trace on 30 June 1958. (c) Illustrated projection of geo magnetic field direction on to vertical plane perpendicular to frontal orientation of irregularity in (b).

found suitable for a determination of speed and irregularity cross-section by the same methods used for the diverging satellite on 30 June 1958. Fig. 17 illustrates a somewhat similar irregularity for this aspect of the phenomenon, and indicates the reflecting surface responsible for the satellite. Here the perpendicular to the front is in the direction 325° and the slope of the irregularity is 22°. Whereas the converging satellite merged with the main trace at 19.40 hr, the diverging satellite, produced by the other face of the irregularity, was not resolved from the main trace until 21.00 hr.

In only five other cases was all the necessary information available for a similar procedure to be adopted. Fig. 18 shows a plot of slopes (with an estimation of the error in deriving them) against the directions perpendicular to the fronts. The continuous curve in the diagram shows the slopes which would be obtained by projecting the geomagnetic field direction on to a vertical plane drawn perpendicular to the front. The consistency between the experimental values and the curve may be significant.

(N, h) curve analysis was carried out for altogether thirty cases of resolved range-spreading. In all cases, the trough of the kink passed over the station prior to the crest.

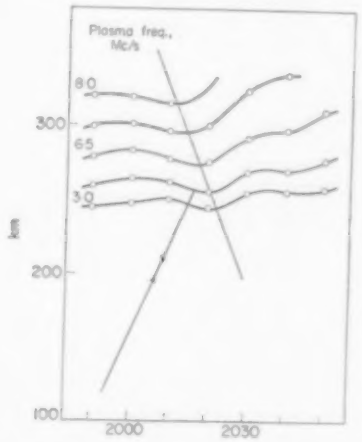


Fig. 17. Variation of contours of equal ionization at the irregularity causing a convergent satellite trace on 9 August 1958.

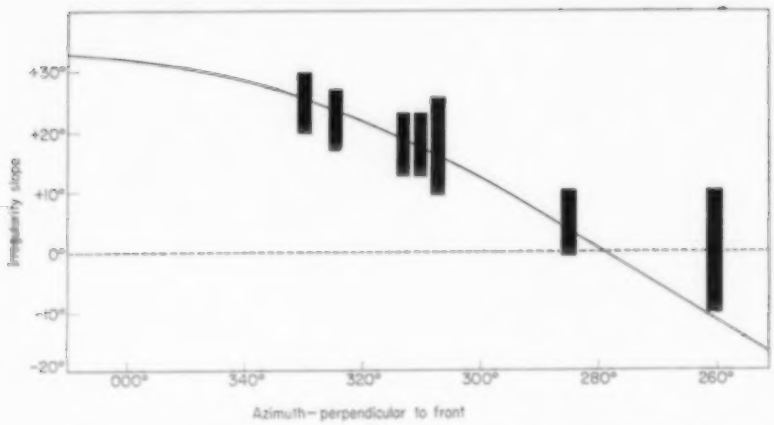


Fig. 18. Plot of irregularity frontal slopes against directions perpendicular to fronts. Solid line represents slopes for projections of geomagnetic field direction on to vertical planes containing azimuths from 010° true to 250° true.

Fig. 19 illustrates the average f_0F2 variation (seventeen cases) for periods either side of the zenithal position of the resolved range-spreading irregularity. The dotted line indicates the average night-time fall in f_0F2 in the absence of irregularities. The difference is roughly a sine curve centred near the zero of time. The maximum departure is of the order of 0·05 Mc/s.

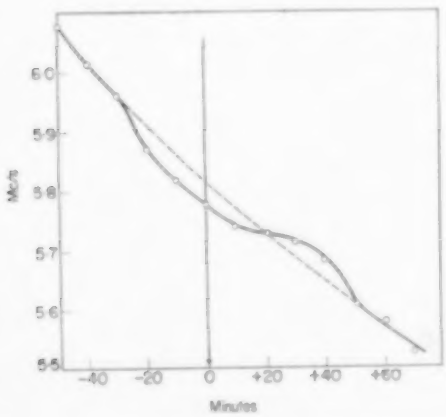


Fig. 19. Variation of average f_0F2 value about overhead position of resolved range-spreading irregularities for seventeen cases.

Fig. 20 depicts a comparison of a vertical section of the irregularities, drawn through the direction of motion, with the estimated plan position of the fronts depicted in Fig. 7. Fig. 20 also shows, for this period, the variations of f_0F2, f_0E_s and of the horizontal geomagnetic field in the direction of the field. The estimated velocity was 175 km/hr in the direction of magnetic north, and the ionization density contours in the meridian plane are drawn to scale. Eight of the ten irregularities recorded have slopes lying close to the direction of the geomagnetic field. This is consistent with the observations on Fig. 18. The positions of these irregularities, determined from (N, h) curves, are consistent with the fronts located by range and direction observations. The f_0F2 value tends to show a dip at the irregularity positions. An E_s cloud is found to be related to one of the irregularities as in Fig. 16. The magnetic record shows substantial variations but these are not related in any simple way to the passage overhead of the ionospheric irregularities.

(ii) *Resolved frequency-spreading.* At 21.30 on 17 December 1944 at Brisbane the ionogram showed a well-defined example of resolved frequency-spreading. This is illustrated in Fig. 15(b). Examination of the electron

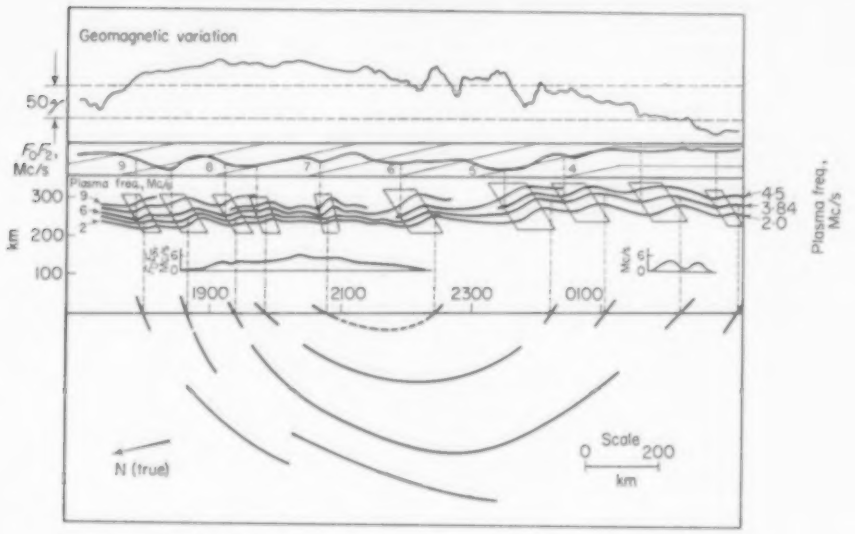


Fig. 20. Variation of contours of equal ionziation density for the night 10–11 July 1958 compared with plan position of irregularities from Fig. 7. Variations of f_0E_s, f_0F2 and horizontal component of geomagnetic field are also shown.

density variation around this time, in the manner described previously, reveals a rippling of the contours similar to that found for resolved range-spreading. The well-defined satellite trace occurrence appeared when the highest point of the contours was vertically above the station (Fig. 21). (This plot is not to scale as the velocity of movement is not known). Also illustrated is the f_0F2 variation at this time. Similar results were obtained in the two other cases investigated.

The variation of f_0F2 around the time of occurrence has been determined for four cases, where the satellite trace is well defined on only one ionogram. The usual occurrence of the phenomenon shows the satellite traces appearing on a succession of ionograms. In the cases chosen the satellite trace results from an isolated front, simplifying the determination of reflection conditions. The average variation of f_0F2 has the same form as that found for resolved range-spreading, but the dip in f_0F2 is some seven times greater (Fig. 22). The centre of the f_0F2 variation, as plotted, shows a phase shift of the quasi-sinusoidal variation when compared with the corresponding variation for resolved range-spreading.

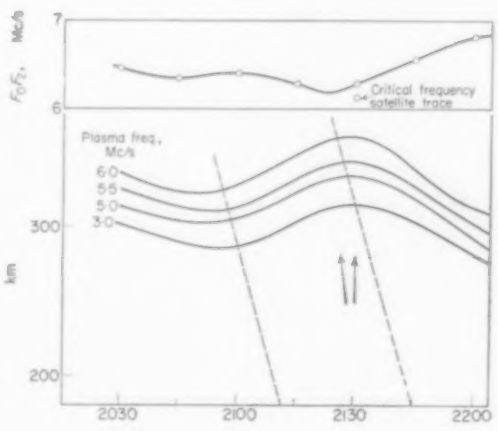


Fig. 21. Variation of contours of equal ionization density at the irregularity causing resolved frequency-spreading at 21.30 on 17 December 1944. The f_0F2 variation is also shown.

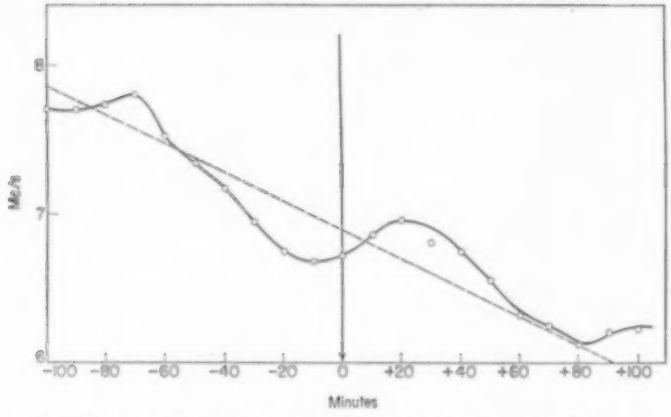


Fig. 22. Variation of average f_0F2 value about overhead position of resolved frequency-spreading irregularities for four cases.

(c) *Other features of irregularities*

(i) *Two-hop unresolved range-spreading.* The fine-structure type of "Spread-*F*", described as unresolved range-spreading, sometimes occurs on the two-hop trace, with no sign of it on the one-hop trace. Fig. 14(a(i)) is an example of this on rotating-loop normal-incidence records and Fig. 14(a(ii)) an example at oblique-incidence. The null at normal-incidence indicates a definite azimuth-of-arrival for these signals. Three-hop traces were not recorded on the normal-incidence films, but at times, at oblique incidence, the spreading of the three-hop trace was more pronounced than that of the two-hop. Fig. 23 shows the result of an investigation of the percentage occurrence, for the hours recorded, of (a) one-hop spreading (whether it occurred on the two-hop or not), and (b) two-hop spreading with no occurrence on the one-hop. Fig. 23 indicates that, while the one-hop occurrence rate is falling, the two-hop occurrence rate first rises, reaches a maximum and falls reaching zero some 15 days after one-hop.

(ii) *Direction-of-arrival of main echo at oblique-incidence.* At times the direction-of-arrival of the oblique-incidence main echo is observed to oscillate about the direction of Armidale. Fig. 24 shows an example, oscillating with a period of about half an hour for 2½ hr. It can be shown that the maximum divergence (15°) only requires a tilt of the ionization contours of about 8°. Such a slope would not be sufficient to produce detectable range-spreading, but Fig. 11(b) illustrates that, about half an hour previous to the commencement of oscillation, range-spreading fronts existed aligned roughly along the Brisbane–Armidale direction. Twenty-one cases occurred on the directional records, where the oscillation

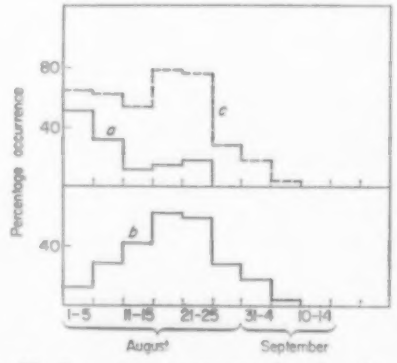


Fig. 23. Percentage occurrence for the time of recording, for (a) one-hop *F2* unresolved range-spreading whether or not this spreading was present on two-hop *F2* trace; (b) two-hop *F2* unresolved range-spreading with an absence of spreading on the one-hop *F2* trace; (c) a combination of histograms (a) and (b).

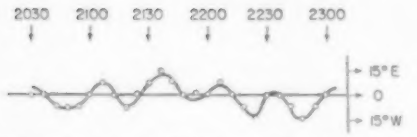


Fig. 24. Sample oscillation of the direction-of-arrival of the main F2 trace from a distant transmitter at Armidale.

lasted for more than an hour, and the period was greater than half an hour.

Figure 25 is a histogram showing the distribution of the main trace direction-of-arrival, during the period of recording, averaging over 6-min intervals, plotted in units of $7\frac{1}{2}°$. There is a pronounced preference for westerly deviations.

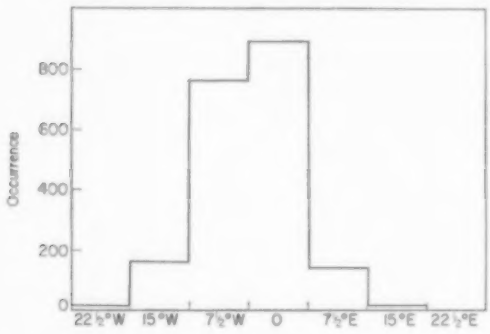


Fig. 25. Distribution of arrival directions for main F2 trace from a distant transmitter at Armidale.

(iii) *Enhanced reflection of satellite traces.* Sometimes the signal strength of the satellite trace of resolved range-spreading becomes enhanced, usually at the time when its range in excess of that of the main trace is approximately 50 km. Curve *d* of Fig. 13 of Ref. (1) shows the intensity variation for one of these traces and Fig. 9(b) shows a swept-gain record. Forty-two cases were observed for 30 nights in April, May and June 1955 when swept-gain records were available. These represent approximately one third of all satellites recorded. Direction-of-arrival records[4], taken at the same time, gave the average elevation of these signals as 59°. The distribution of azimuths of these forty-two cases was very similar to that found when all satellites are considered (cf. Fig. 19 of Ref. (1) and Fig. 8 of this paper), 79 per cent of cases having their azimuth in the NW quadrant.

If there is enhanced reflection for signals with a particular excess range, in the case of unresolved range-spreading, ionograms with spreading widths of this excess range should occur more often than expected, at the expense of smaller widths. Fig. 26(a) shows a histogram of spreading widths for 2185 ionograms during June and July 1952 at Brisbane. An interruption to the systematic fall with increased range width occurs at 50 km, where there is a slight increase in occurrence. Fig. 26(b) also shows a similar histogram from Townsville ionograms for this period. This indicates a similar effect at the 90 and 100 km position. Ionograms with E_s present had to be discarded for the Townsville histogram due to confusion with the $F+E_s$ echoes at about 100 km excess range. 998 ionograms were used.

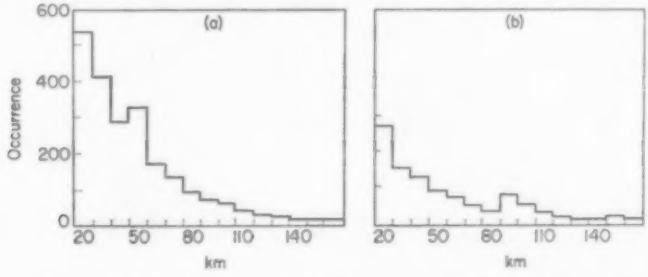


Fig. 26. Distribution of unresolved range-spreading widths for June and July 1952, (a) for Brisbane (dip. 57°); (b) for Townsville (dip. 47°).

(d) *Statistics from ionograms*

(i) *Introduction.* The previous results of this paper appear to indicate that all four classes of spreading defined in Section 1 result from the same general type of irregularity structure, even though the dimension of the structures differ considerably. This allows a grouping of these classes for a determination of diurnal, seasonal and sunspot cycle variations of the phenomenon.

(ii) *Diurnal variation.* Range-spreading and frequency-spreading show very different diurnal distributions. Fig. 27(a) and 27(b) show the results for Brisbane for the winter and summer months of 1952 compared with similar results for Townsville (lat. 19·3°S., long. 146·7°E., magnetic dip 47°) and Hobart (lat. 42·9°S., long. 147·2°E., magnetic dip 72°). The Townsville and Brisbane occurrences are comparable but for Hobart there is a marked increase in spreading which also extends into the daylight hours. Only 8 per cent of all spreading at Hobart is range-spreading whereas this figure is 61 per cent for Brisbane and 64 per cent for Townsville. At Hobart only 5 per cent of all spreading is resolved, while 27 per cent is resolved at Brisbane and 38 per cent at Townsville. For range-spreading Fig. 27(a) shows the average *h'F2* variation for the period concerned, while Fig. 27(b) shows the average f_0F2 variation for frequency-spreading, plotted negatively. The *h'F2* values represent the lower boundaries of the *F2*-layer.

For Brisbane and Townsville range-spreading has roughly a peak of occurrence at midnight falling to zero in the sunset and sunrise periods. This simple distribution seems to be modulated by a tendency for increased occurrence at the higher values of *h'F2*. Frequency-spreading has a peak just before sunrise, and seems to have some relationship with critical frequency (f_0F2).

As previous work by Singleton[2] had shown that there is a strong inverse correlation

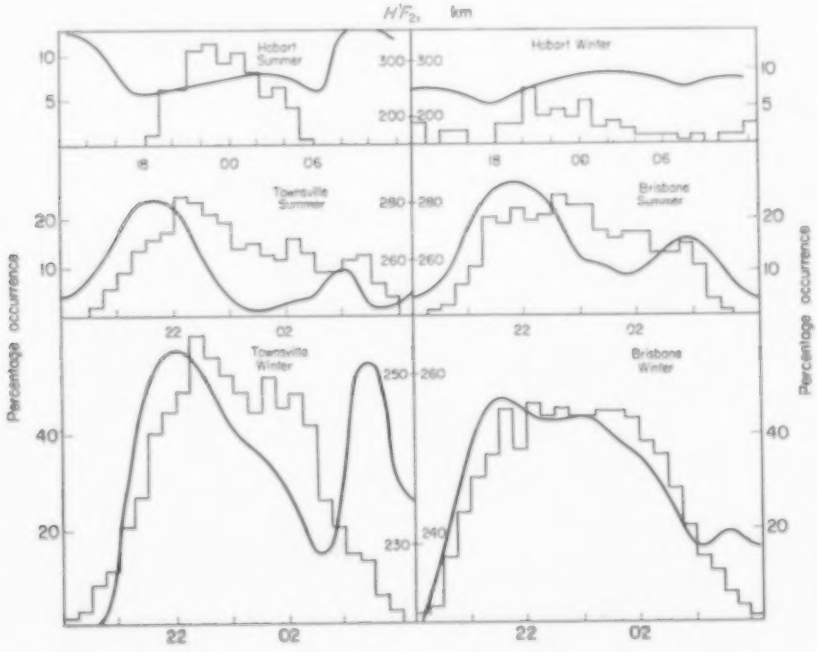


Fig. 27(a). Diurnal distributions: Range-spreading at Hobart, Brisbane and Townsville for summer and winter, 1952, with average *h'F2* variations.

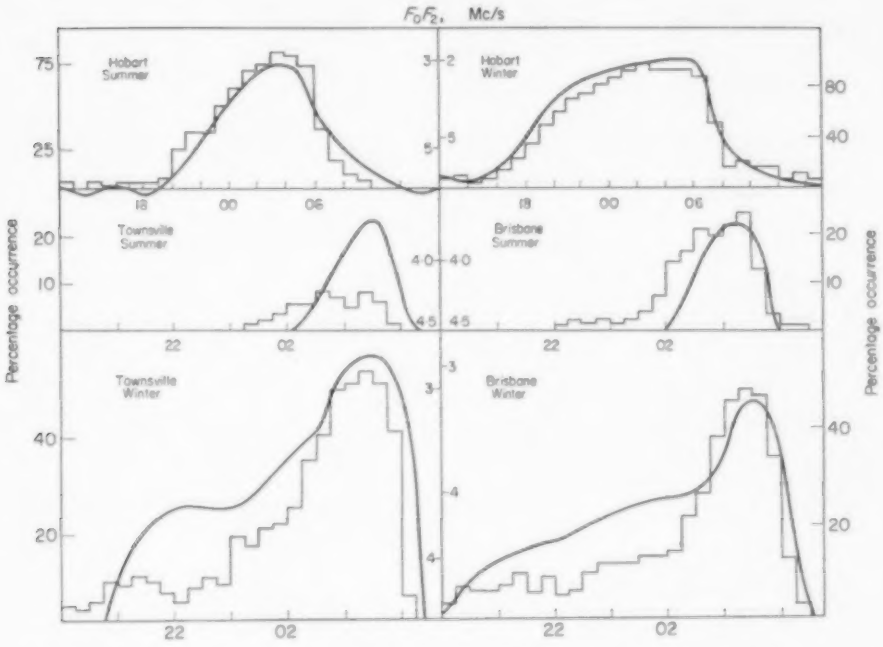


Fig. 27(b). Frequency-spreading at Hobart, Brisbane and Townsville for summer and winter, 1952, with average f_0F2 variations plotted negatively.

between f_0F2 and frequency of occurrence of frequency-spreading, an endeavour has been made to detect any additional diurnal effect by taking results only for months in which the average (night) f_0F2 fell between 3 and 4 Mc/s and considering only ionograms in which f_0F2 lay between 3·1 Mc/s to 3·5 Mc/s inclusive. Approximately 2000 such ionograms were obtained between June 1943 and August 1955. The results are shown in Fig. 27(d). The frequency-spreading distribution is considerably different from that of Fig. 27(b) revealing a second maximum. It is evident that there are pronounced diurnal changes in frequency of

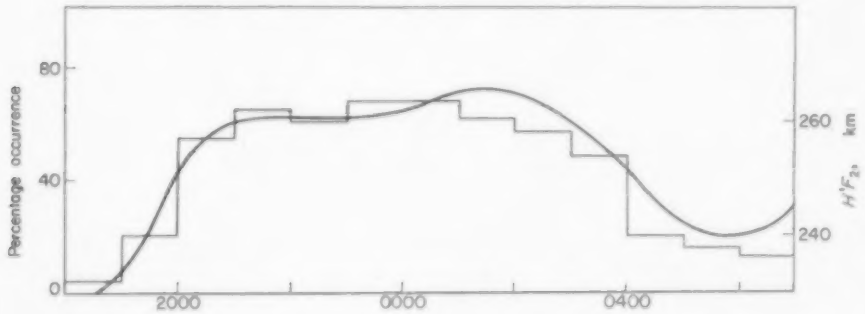


Fig. 27(c). Range-spreading (Brisbane) for ionograms having approximately equal f_0F2; with average $h'F2$ variations.

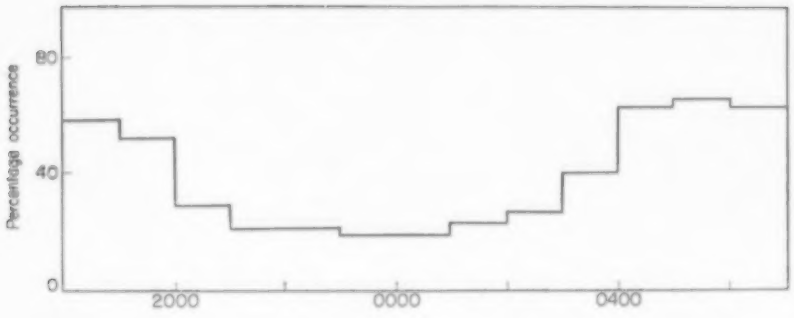


Fig. 27(d). Frequency-spreading (Brisbane) for ionograms having approximately equal f_0F2.

occurrence apart from those associated with f_0F2 changes. Fig. 27(c) shows the histogram for range-spreading from the same records compared with the average $h'F2$ for the months chosen.

(iii) *Seasonal variation.* Fifteen years of ionograms have been analysed (1944 to 1958 inclusive). The results for range-spreading and frequency-spreading combined are shown in Fig. 28. This was done because their separate distributions were not significantly different. Curve *A* represents the distribution for years when the average sunspot number was less than 100 while Curve *B* is the distribution for years when the average sunspot number was greater than 100. These curves show a maximum in winter, and a sub-maximum in summer, with minima around the equinoctial periods. It is interesting to note that the distributions are not symmetrical about the solstices and the equinoxes, and the minima are displaced about a month from the equinoxes in the direction of the summer solstice.

(iv) *Sunspot-cycle variation.* The sunspot cycle variation of "Spread-*F*" in half-yearly periods (range-spreading and frequency-spreading combined) is shown in Fig. 29. This indicates a decided inverse correlation with sunspot activity. However, peaks of spreading occurrence are slightly displaced from the

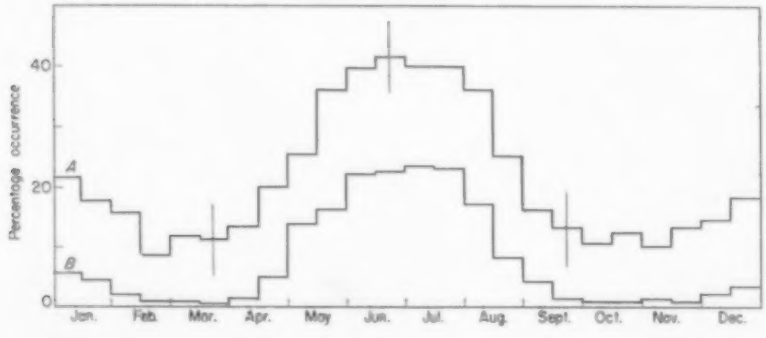


Fig. 28. Seasonal variation of "Spread-*F*". Curve *A* represents years (9) when average sunspot number was less than 100. Curve *B* represents years (6) when average sunspot number was greater than 100.

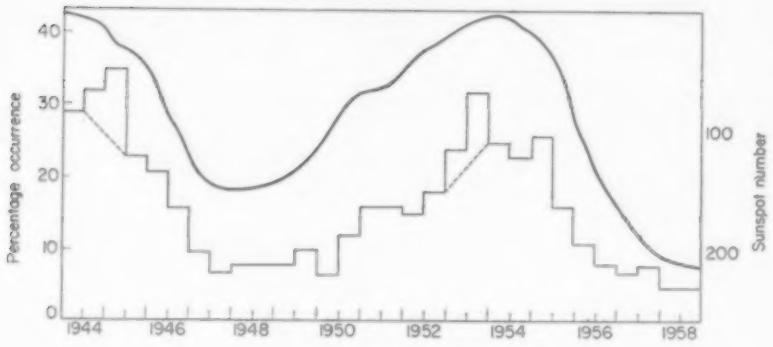


Fig. 29. Sunspot cycle variation of "Spread-*F*". Sunspot number variation shown plotted negatively.

minima of sunspot occurrence. It should be noted that, although spreading on one-hop *F*-traces is almost exclusively a night-time phenomenon at Brisbane, the percentage occurrences for Fig. 28 and Fig. 29 refer to the full 24 hr.

Acknowledgments—The author is indebted to Professor H. C. Webster for his interest and guidance in this work and would like to acknowledge the generous help, at all times, from the academic and technical staffs of the Physics Department, particularly Dr. R. W. E. McNicol and Dr. J. A. Thomas who gave invaluable assistance. Thanks are also due to Mr. M. J. Burke, Mr. H. MacGregor, Mr. J. B. Mott and Mr. K. Perry for their contributions to the project. The cooperation of the University of New England Armidale, is gratefully acknowledged. The Ionospheric Prediction Service kindly supplied the ionograms used in this analysis. This work was financially supported by the Radio Research Board of C.S.I.R.O. Most of the work was carried out while the author was the holder of a C.S.I.R.O. Studentship.

REFERENCES

1. R. W. E. McNicol, H. C. Webster and G. G. Bowman, *Aust. J. Phys.* **9**, 247 (1956).
2. D. G. Singleton, *Aust. J. Phys.* **10**, 60 (1957).
3. W. Ross, *J. Instn. Elect. Engrs.* **94**, (part 3), 99 (1947).
4. J. A. Thomas and R. W. E. McNicol, *Proc. Instn. Elect. Engrs.* **102B**, 793 (1955).
5. E. R. Schmerling, *J. Atmos. Terr. Phys.* **12**, 8 (1958).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 150–156. Printed in Great Britain.

FURTHER STUDIES OF "SPREAD-*F*" AT BRISBANE–II

INTERPRETATION

G. G. BOWMAN

Physics Department, University of Queensland, Brisbane

(*Received* 10 *September* 1959)

Abstract—The experimental results for the phenomenon, revealed from the analysis presented in Part I of this paper, are discussed. It is deduced that "Spread-*F*" results from a ripple structure in the ionization contours of the *F*2-layer. Model ionization distributions are proposed, and ray-paths drawn, in an endeavour to explain the various aspects of the types of "Spread-*F*" classified in Part I. The evidence suggests that the type of spreading observed is determined by, (i) the f_0F2 variation at the time, (ii) the amplitude of the ripples, and (iii) the spacing between the ripples. If the ripple amplitude is small, no spreading will be observed. The ripples appear to be present all the time, and their amplitudes seem to become greater with an increase in *F*2-layer height.

1. INTRODUCTION

This paper consists of a discussion of experimental evidence which is presented in the preceding paper (herein referred to as Part I). A summary of the results of analysis from Part I is contained in Table 1.

The various types of "Spread-*F*", classified in Part I, are discussed in turn, with models of ionization contours, which are proposed to explain the observed aspects of the phenomena.

Table 1.

Type of spreading ⟶	Resolved range	Unresolved range	Resolved frequency	Unresolved frequency
Plan position				Frontal
Vertical section				
Frontal spacings	Greater than 60 km	20–60 km		Discrete echoes
f_0F_2 variation				

2. DISCUSSION

(a) *Resolved range-spreading*

The present observations (consistently with those reported by McNicol *et al.*[11]) indicate that this phenomenon is associated with a front or fronts having a linear extent of the order of 1000 km, and where several fronts are present, the spacing between them is of the order of 100 km. Vertical sections, perpendicular to the front, show a kinking of the contours of equal ionization density, corresponding points of the kink being progressively displaced with increasing height (Fig. 16 of Part I). The lines joining corresponding points of the successive kinks, appear to be parallel to the projection of the geomagnetic field direction on the vertical plane perpendicular to the front. The effect of the kinking is to produce approximately parallel ionization contours at an angle to the horizontal. Fig. 19 of Ref. (1) shows that the fronts with normals lying in the NW–SE quadrants (i.e., the majority of cases) appear to travel in directions between N and W. This suggests that in these cases the crest lies, on the average to the SE of the trough, and reflection leading to a satellite trace is visualized as occurring as illustrated in Fig. 1(a).

The fact that the rotating-loop records showed only a few cases of the front approaching the station rather than receding, is well explained if it is assumed that all irregularities in question were of the character illustrated in Fig. 1(a), for this would reflect only after the irregularity had passed over the station. The marked preference for the main trace from Armidale to wander towards the west rather than the east (Fig. 25 of Part I) is also explained by a kinking of this kind in the ionization contours, for during the period of the rotating-loop observations nearly all fronts recorded had their normals in the NW quad-

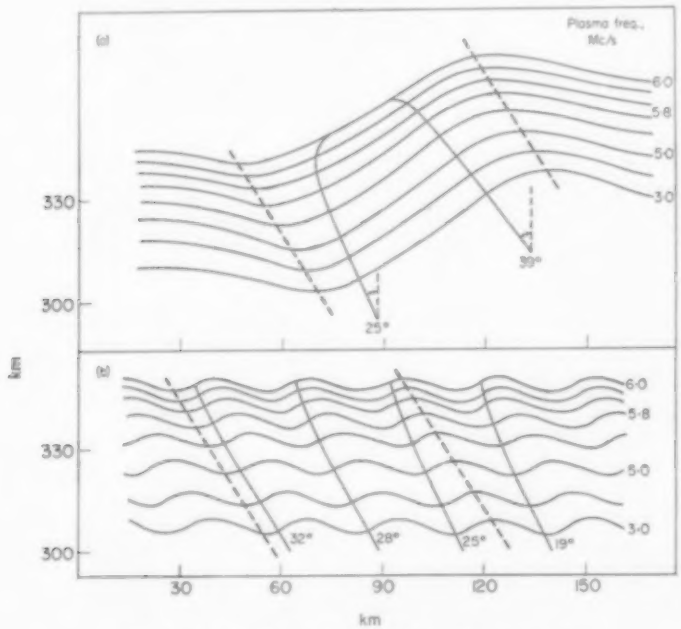


Fig. 1. (a) Model of ionization distribution at a resolved range-spreading irregularity. (b) Model of ionization distribution at unresolved range-spreading irregularities.

rant (Fig. 8 of Part I). Background height changes of the layer will increase an upward tilt in the ionization contours when the layer is rising, and accentuate a downward tilt when the layer is falling. This will explain the occurrence of diverging satellites when the layer is rising and converging satellites when the layer is falling (cf. Fig. 25 of Ref. (1)). The number of divergent satellites exceeds those of the convergent type (Fig. 23 of Ref. (1)) because the background height change has to be much greater to produce a convergent type than that required for a divergent type, for any irregularity similar in shape to the one illustrated in Fig. 1(a). Occasionally the downward slope will be pronounced enough to record the few convergent type satellites (Fig. 17 of Part I) observed during the period of recording by the rotating-loops (a period when the height of the layer is generally rising).

In Fig. 1(a) ray-paths are drawn for the *o*-ray (for a frequency close to f_oF2) in a model ionization distribution assuming the perpendicular to the front to lie in the magnetic meridian. The diagram illustrates that reflection is possible at vertical angles of 25° and 39°, at positions where the changing orientations of the ionization contours is such that only a narrow beam will be reflected. In the interval between the occurrence of these two reflection conditions, the position of the frontal irregularity moves approximately 100 km. At some point between these positions the reflection conditions encountered would be very similar to those for vertical incidence, and a much stronger signal can be expected than that at 25° or 39° vertical angle. A stronger signal at this point could possibly explain the enhanced signals reported in Section 3(c) of Part I.

From a study of the (N, h) curves, taken in conjunction with the f_oF2 values, it seems possible that the kinking arises from the displacement of ions and electrons along the lines of magnetic force, leading to a rarefaction in the trough region, and a concentration in the crest region at the positions of maximum ionization.

The movement of the configuration need not necessarily be perpendicular to the front (cf. Fig. 7 of Part I). In some cases, at least, the configuration is extremely stable over a period of some hours, as suggested in Ref. (1).

(b) *Unresolved range-spreading*

This type of spreading is associated with fronts of a similar nature to those responsible for resolved range-spreading. Frontal spacings range from 20 to 60 km having a peak of occurrence at 35 km (Fig. 10 of Part I). The distribution of normals is somewhat similar to that for resolved range-spreading, but is shifted some 30° towards the north. Fig. 1(b) illustrates the suggested vertical section contours and *o*-ray reflection paths for this phenomenon, using another model of ionization distribution. The illustrated cross-section lies in the plane containing the magnetic meridian. Here the drawn ray-paths originate at the one point on the ground.

In general, the movement of the configuration will not be perpendicular to the front, so that as the pattern moves with respect to the recording station, the reflecting point, on each front, will move along the line of the front as well as around the curvature of the vertical section of the front. Any irregularities encountered in the vertical sections or the frontal curvatures will cause fading and may possibly account for the diffuse nature of the trace when integrated over periods of a minute or more (Fig. 9a, ii of Part I).

The pulse width on the rotating-loop records corresponds to approximately 10 km. Slant reflections with 260 km range, when the height of the reflection is 250 km, will come from a region displaced approximately 16° from the zenith. Irregularities causing tilts of less than 16° will therefore not give satellites on the one-hop trace, but may possibly do so on the two-hop in the manner illustrated in Fig. 2. During the sunspot maximum period, between the winter solstice and the equinox, one-hop spreading occurrence falls to nearly zero (Fig. 28 of Part I). If the average ripple amplitude were to gradually diminish between the winter

solstice and the equinox we should expect the frequency of occurrence of one-hop spreading to fall to zero before the frequency of occurrence of two-hop spreading. At an intermediate stage, the occurrence of two-hop spreading might well show a maximum, as depicted in Fig. 23 of Part I. (The possibility of the two-hop spreading being due to ground scattering appears to be ruled out by the resolution of satellite traces, and the presence of nulls, on rotating-loop records). It will be noted that the structures envisaged by Baird[2] to explain high-multiple reflections are generally similar to those suggested here, save that the ratio of amplitude to "wavelength" is much less.

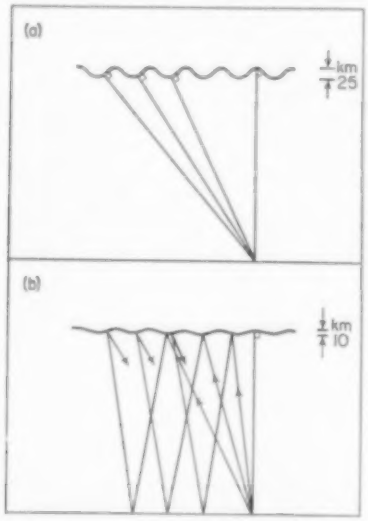


Fig. 2. (a) Ionization distribution with ripple amplitude sufficiently large to give one-hop spreading. (b) Ripple amplitude much smaller than that in (a). Reflection paths for two-hop spreading illustrated.

Enhanced echo strength sometimes occurs for this phenomenon, at about 50 km excess of vertical-incidence range at Brisbane, and 90–100 km excess range at Townsville (Fig. 26 of Part I). These distances are approximately the excess ranges to the *F2*-layer, along the lines of force of the earth's magnetic field. Since the reflected signals, at least for Brisbane, have azimuths of reflection reasonably close to magnetic north (Fig. 12 of Part I), it suggests that the geomagnetic field orientation is, in some way, responsible for contours which sometimes give enhanced echoes for signals having the same direction as the geomagnetic field.

(c) *Resolved frequency-spreading*

The distribution of reflection azimuths in the NW and SE quadrants for this type of spreading is consistent with the irregularity being frontal in character, the frontal orientations being similar to those for resolved range-spreading. Assuming the observable movement to be NW, the number of points in the SE quadrant suggest that satellite traces will be recorded as the irregularity approaches the station. An approximately equal number of points is located in the NW quadrant representing satellite traces recorded as the irregularity moves away. Since the measured elevations-of-arrival of the satellite echoes (at 2·28 Mc/s) were around 87°, these echoes must come from parts of the ionosphere close to the zenith. The vertical sections of the irregularities associated with this phenomenon are essentially similar to those for resolved range-spreading, the satellite echo being most pronounced when the recording station is immediately under the crest of the irregularity. The variation of f_0F2 about the irregularity position is the same nature as that for resolved range-spreading (Fig. 22 of Part I), except that departure from the normal is much greater, being of the order of 0·5 Mc/s. Also, for frequency-spreading, the drop in f_0F2 between points say one hour either side of the central plotting position is twice as great as the drop over the similar period for range-spreading.

In Fig. 3 ray-paths have been plotted for a fixed frequency (near the f_0F2 value) for the *o*-ray in the presence of the geomagnetic field. Ionisation contours similar to those shown in Fig. 21 of Part I were used for a model and the f_0F2 variation was made approximately the same as indicated in this figure. The normal to the front lies in the magnetic meridian plane.

Two forms of the irregularity are considered. In Fig. 3(b) the curvature at the crest of the irregularity is pronounced, the radius of curvature being somewhat less than the height of the base of the layer. In this case results show that only one ray-path will be returned to the sender by reflections in the vicinity of the crest. In Fig. 3(a) the radius of curvature is made greater than the height of the base of the layer. Here there are two ray-paths, reflected from points about 80 km apart near the crest of the irregularity, which return to the sender. The ray-paths on each diagram of Fig. 3 meet at the ground.

This suggests that resolved frequency-spreading occurs when the curvature of the crest of the ripple is small enough to allow two reflecting paths near the zenith, plus the added condition that the difference between the f_0F2 values, at the two reflection points, is sufficiently large for the two traces to be distinguishable from each other on an ionogram. On the other hand, with a larger curvature range-spreading becomes possible

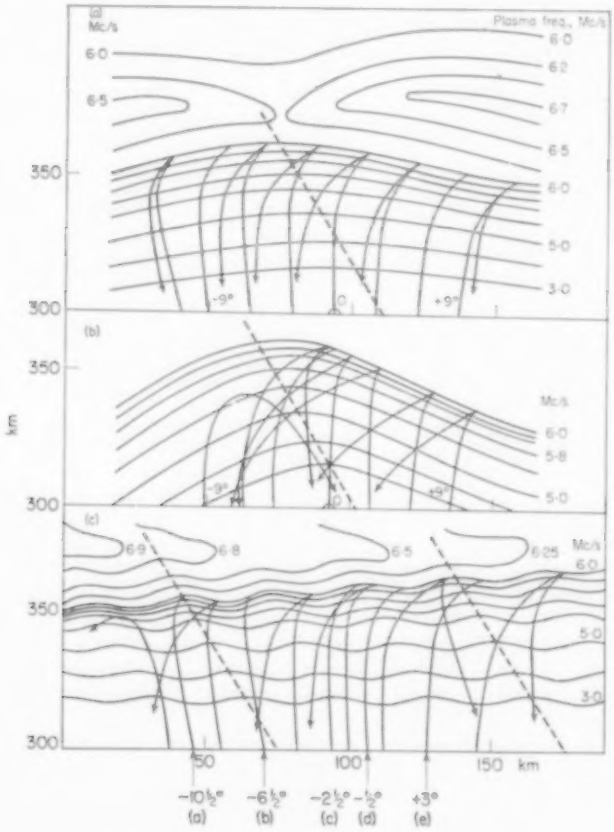


Fig. 3. (a) Model of ionization distribution for resolved frequency-spreading with ray-paths for o-ray. Contours in plane of magnetic meridian. (b) Similar model to (a) with much greater curvature in ionization contours. (c) Model of ionization distribution for unresolved frequency-spreading.

(cf. Fig. 1a). This would explain the complementary character of the diurnal distribution of range-spreading and frequency-spreading (Fig. 27 of Part I). The difference of approximately 15 min between the centre of the characteristic f_0F2 variations, as between resolved range-spreading and resolved frequency-spreading, is consistent with the upward slope of the ionization contours being responsible for range-satellites and the crest of the contours (which follows the upward slope) being responsible for frequency-satellites.

By taking the integral $\int \frac{1}{u}\,ds$ (an approximate expression for group path) along the two ray-paths of Fig. 3(a), which are returned to the sender, the virtual range (for a 6 Mc/s signal) for the path in the region where f_0F2 is 6·2 Mc/s was found to be 435 km, while a virtual range of 396 km was calculated for the path where f_0F2 is 6·7 Mc/s. These calculations are consistent with the observed features of resolved frequency-spreading.

(d) *Unresolved frequency-spreading*

On the few examples of this type of "Spread-F" recorded by the rotating-loop system, records show resolution of traces with a specific direction-of-arrival indicated, the records looking very similar to those of unresolved range-spreading. However, here the spacing between traces represents a change in group-retardation of the signal, whereas for unresolved range-spreading the spacing represents, in most cases, solely a change of range for the reflected signal.

Figure 3(c) illustrates the possible reflection paths for the *o*-ray for this phenomenon, assuming ripples of small amplitudes, having frontal spacings of 30 km and a f_0F2 gradient across the ripple structure. Again the cross-section lies in the plane of the magnetic meridian. As mentioned in Section 2(b), for a base height of layer of 250 km, echoes with a zenith angle of arrival of less than 16° will not appear as range-spreading. There is a circle of diameter 143 km from within which the excess range of a satellite signal cannot easily be detected. Fig. 3(c) shows that within this circle, for the model chosen, five reflection paths return to the sender. The values of the integral $\int \frac{1}{u}\,ds$ for the ray-paths marked *a*, *b*, *c*, *d* and *e* were found to be 402, 405, 420, 417 and 424 km, respectively. Also, the f_0F2 value changes from 6·8 Mc/s to 6·3 Mc/s between the outside members of this group of ray-paths. These calculations are consistent with the observed features of unresolved frequency-spreading.

(e) *General*

Bramley and Ross[3] found rapidly changing lateral deviations, with quasi-periods of the order of seconds, for the direction-of-arrival of main echoes originating from a distant transmitter. If these occurred during the present investigation they could not be followed with the recording equipment used. In addition, Bramley and Ross reported a slowly varying component of deviation with quasi-periods up to half an hour or more. These have been detected by the Armidale–Brisbane experiment (Fig. 24 of Part I). This investigation has also recorded echoes additional to the main trace, and these have been termed satellite traces. These show lateral deviations through angles as large as 90°. The experimental analysis of this investigation suggests a rippling of the ionization contours of the *F2*-region. This is in line with the wrinkled ionosphere proposed by Bramley and Ross to explain their tilts. The oscillation of the direction-of-arrival of the main trace can be followed across the ripple structure particularly when the linear aspect of the structure is oriented along the transmission path and the ripple amplitude is small. At mid-latitudes, at normal-incidence, resolved satellite traces have a peak of occurrence[4], and this appears to originate from large scale ripples with amplitudes large enough to give the satellite traces recorded at Brisbane from a transmitter 355 km away.

Even though the diurnal-variation histograms (percentage occurrence independent of f_0F2 variation) represent spreading occurring when the f_0F2 value lies between 3·1 and 3·5

Mc/s, it may be assumed that similar distributions will result at other frequencies. If these histograms (range-spreading and frequency-spreading) are combined (Fig. 4), the resulting histogram suggests that ripples responsible for "Spread-*F*" are present all the time, sometimes being responsible for range-spreading records and at other times for frequency-spreading records. If the ripple amplitude is very low no "Spread-*F*" records will result. Frequency-spreading would be seen when the amplitude of the ripple was relatively shallow and the f_0F2 variation gradient was appreciable.

observations contained here, are the hydromagnetic-resonance wave theory[5] or the electron-resonance wave theory[6] of Hines, or the cellular wave theory of Martyn[7]. The variation of the disturbance slope with frontal orientation is approximately that expected from Martyn's theory. Movement of the irregularity structure in a direction other than perpendicular to the frontal orientation (Fig. 7 of Part I) may be an obstacle for the theories of Hines. A complete theory would also need to explain the presence of E_s-clouds along the line of slope of the kinks extended to the 100 km level.

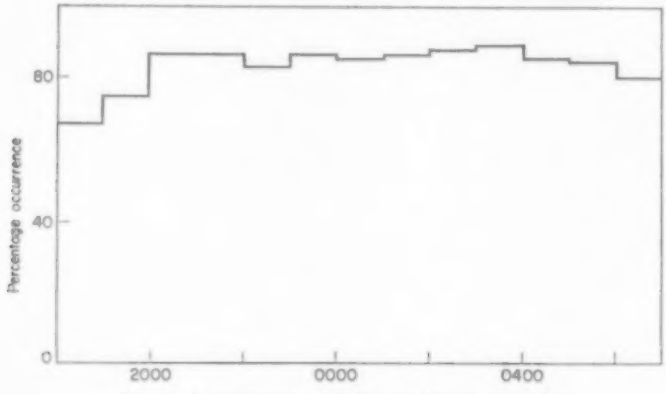


Fig. 4. A combination of diurnal distributions for range-spreading and frequency-spreading independent of f_0F2-factor. (Fig. 27 (c), (d) of Part I).

If we postulate that, for Brisbane, the ripple amplitude increases with an increase in the height of the layer, both conditions for frequency-spreading are satisfied around the sunset and sunrise periods when the phenomenon has peaks of occurrence (Fig. 27(d) of Part I). Here $h'F2$ is usually a minimum and df_0F2/dt a maximum. When the occurrence of range-spreading is compared with the average $h'F2$ variation (Fig. 27(c) of Part I), it also seems likely that the ripple amplitude will be more pronounced at greater heights. Fig. 15(c) of Part I appears to support this hypothesis.

(f) *Origin of ripples*

Theories which could possibly account for the

Acknowledgments—The author would like to thank Professor H. C. Webster and Dr. J. A. Thomas for helpful discussion during the preparation of this paper. The work was financially supported by the Radio Research Board of C.S.I.R.O., and was carried out while the author was the holder of a C.S.I.R.O. Studentship.

REFFERENCES

1. R. W. E. McNicol, H. C. Webster and G. G. Bowman, *Aust. J. Phys.* **9**, 247 (1956).
2. K. Baird, *Aust. J. Phys.* **7**, 165 (1954).
3. E. N. Bramley and W. Ross, *Proc. Roy. Soc.* A**207**, 251 (1951).
4. R. W. E. McNicol and G. G. Bowman, *Aust. J. Phys.* **10**, 588 (1957).
5. C. O. Hines, *J. Atmos. Terr. Phys.* **7**, 14 (1955).
6. C. O. Hines, *J. Atmos. Terr. Phys.* **9**, 56 (1956).
7. D. F. Martyn, *Proc. Roy. Soc.* **201**, 216 (1950).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 157–164. Printed in Great Britain.

THE ORIGIN OF NITROGEN IONIZATION IN THE UPPER ATMOSPHERE*

J. W. CHAMBERLAIN and C. SAGAN

Yerkes Observatory, University of Chicago, Williams Bay, Wisconsin

(*Received* 13 *November* 1959)

Abstract—Although the N_2^+ First Negative bands in twilight are probably the result of resonance scattering by N_2^+ ions, the symmetry of morning and evening observations make it unlikely that the ionization is produced by sunlight in the *E*-region. If the flux in the He II λ304 line is as high as 15 erg cm^{-2} sec^{-1}, the observed First Negative system intensities can be accounted for, and the nitrogen ionization is localized in the *F*-region. If the flux in the λ304 line is much smaller, the ionization must be produced by incident charged particles in the *E*-region.

1. INTRODUCTION

Slipher[1] first reported a twilight enhancement of the airglow, the First Negative bands of N_2^+ $B^2\Sigma_u^+ \longrightarrow X^2\Sigma_g^+$. Later observations were made by Elvey[2]. Some 8 or 9 bands may be recorded photographically, and Dufay[3] has estimated the relative intensities of λ3914 (0-0), λ3884 (1-1), λ4278 (0-1) and λ4737 (1-2). That the intensity of the emission was actually variable from night to night was demonstrated by J. and M. Dufay[4]. They found a correlation of the enhancement with magnetic activity and with the occurrence of aurorae far to their north in the auroral zone.

Bates[5] has shown that the excitation process for twilight N_2^+ emission is the Wulf–Deming[6] mechanism of resonance scattering,

$$N_2^+(X^2\Sigma_g^+) + h\nu_{\lambda 3914} \longrightarrow N_2^+(B^2\Sigma_u^+). \quad (1)$$

Measurement of the intensities of scattered light can give important information on the N_2^+ ionic densities in the upper atmosphere and provides a technique of ionospheric analysis that has so far hardly been initiated. An accurate photometric study of the emission in the First Negative system has not yet been performed; the rough intensity estimates that can be made are discussed in Section 2. In the following pages we discuss possible mechanisms for producing the ionized nitrogen required for reaction (1).

2. OBSERVATIONAL RESULTS

In resonance scattering the photon intensity $\mathcal{J}(\mu)$, leaving the scattering layer in direction θ ($\mu = \cos\theta$) measured from the normal to the layer, may be written

$$4\pi\mu\,\mathcal{J}(\mu) = g\,\mathcal{N}'(z_s). \quad (2)$$

Here, $\mathcal{N}'(z_s)$ is the effective number of fully illuminated N_2^+ molecules above z_s, the height of the solid-earth shadow in the direction of observation, and g is the number of photons scattered per second by a fully illuminated ion, which is a quantity depending on the solar flux and the *f*-value[7]. The quantity g depends on whether the N_2^+ ions in the $X^2\Sigma_g^+$ state are in vibrational equilibrium with the kinetic temperature. If the gas is excited by a large flux at extremely low pressures, a molecule may remain in the ground electronic state for a time short compared with the time between vibrational collisions. In this case the higher

* The research reported in this paper was supported in part by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19(604)-3044 with the University of Chicago.

vibrational levels could become appreciably populated. However, we shall take essentially all the N_2^+ as being in $X^2\Sigma_g^+$, $v''=0$; then Bates'[5] computation, modified with Shull's[8] improved f-value, gives for the 0–0 band, λ3914, $g=0{\cdot}20$ photon sec^{-1} ion^{-1}.

Some rough estimates of the λ3914 twilight intensities have been made by Bates[5, 9] from qualitative results of Dufay and Dufay[4]. Bates gives the estimated intensities along the direction of observation (zenith angle ζ) for an apparent height z_a. If the decay of the twilight-glow is governed by ionizing radiation in the far ultraviolet, with a very large screening height, z_s is no longer a convenient independent variable. Therefore, we have translated these heights to an observer's angle of solar depression, α (see Fig. 1), by means of the

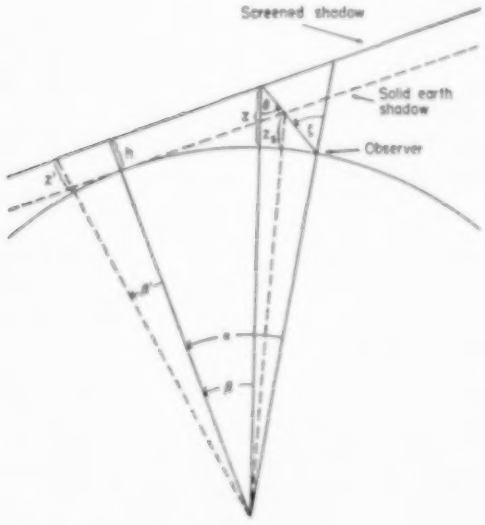


Fig. 1. Geometry of the problem.

expression, valid for observations in the vertical plane through the sun,

$$\frac{z_a}{a} \simeq \frac{1}{2}\,\frac{[\sin\zeta - \sin(\zeta-\alpha)]^2}{\cos^2(\zeta-\alpha)}, \qquad (3)$$

where a is the Earth's radius. With observations at $\zeta=70°$ and $\zeta=80°$, Bates'[9] estimates in rayleighs are that $4\pi\mathcal{J}$ is normally less than 100 R for $z_s=100$ km ($\alpha=13°$), and always less than 10 R for $z_s=200$ km ($\alpha=20°$). [A rayleigh, R, is equivalent to 10^6 photons cm^{-2} (column)$^{-1}$ sec^{-1}.] In referring these intensities to the zenith, they should be diminished by a factor of about 3, depending somewhat on the true emission altitude.

However, these upper limits appear to be too small. Swings and Nicolet[10] found Na and N_2^+ twilight emissions to be of comparable intensity under quiet magnetic conditions on a panchromatic plate in winter, when the zenith Na emission is usually several kilorayleighs. Also, an emission of 10 R is very weak; the upper limit imposed by a failure to record the twilight flash at $z_s \geqslant 200$ km is probably higher than this value. Estimated emission rates for λ3914 and the corresponding N_2^+ abundances may be found in the third and fourth columns of Table 2, which were estimated from measurements near the horizon, but referred to the zenith and corrected for tropospheric extinction. During times of disturbed magnetic activity the intensity increases, and it appears that then the zenith emission may exceed 1 kR for rather low values of z_s (<100 km). Normally the emission is probably within an order of magnitude of the values in Table 2.

3. THE POSSIBILITY THAT N_2^+ IS PRODUCED PRIMARILY IN THE *E*-REGION BY SUNLIGHT

M. Dufay[3] found that the N_2^+ bands were rarely detectable when the solid-earth shadow was at a height $z_s<90$ km or at $z_s>125$ km. The lower limit may have no significance insofar as emission height is concerned, because for low values of z_s, the strong solar continuum completely obscures the emission. Swings and Nicolet[10] could detect λ3914 at much lower apparent heights. The higher altitude (125 km) is probably indicative of a true decrease in the N_2^+ ion density.

The true height of emission, z_e, is not so readily obtained. Dufay was not able to make a precise estimate of the screening height, although a value of 20 km, which would be expected for violet light, did not seem incon-

sistent with his rough comparison of zenith and horizon intensities.

But if we interpret the observations as indicating the bulk of the N_2^+ ions to be within one or two scale heights above a true height of 115 km or so, a difficulty arises. Swings and Nicolet[10] found that the intensity ratio of twilight/dawn emission is approximately unity. If the N_2^+ ions are produced by solar ultraviolet or X-radiation, which has a screening height enormously greater than 20 km, any N_2^+ observed in the morning (before the ionizing radiation reaches the *E*-region) must result from blue light scattered by N_2^+ ions that have existed all night.

Recombination of N_2^+ ions will proceed rapidly through dissociative recombination[11],

$$N_2^+ + e \longrightarrow N^* + N^{**}, \qquad (4)$$

where the nitrogen atoms may be in excited terms. For *E*-region temperatures the experiments of Bialecke and Dougal[12] indicate a rate coefficient for reaction (4) of $\alpha_{rec} = 7 \times 10^{-7}$ cm^3 sec^{-1}. The mean lifetime of N_2^+, when there is no longer any incident ionizing radiation, is $\tau_{rec} \simeq (n_e \alpha_{rec})^{-1}$. If we take $n_e \simeq 10^4$ electrons cm^{-3}, appropriate to early evening, we find $\tau_{rec} \simeq 2$ min.

The conclusion is quite definite. The morning–evening symmetry cannot be explained on the basis of ionization produced in the *E*-region by sunlight. We will see in Section 4.2 that all solar ultraviolet and X-radiation with incident fluxes even approaching that required to explain the observed nitrogen ionization are absorbed far above the *E*-layer in twilight. Thus it is necessary either that the emission actually arise from a much greater altitude, or that some source other than sunlight is responsible for the ionization. We discuss these possibilities in turn.

4. THE POSSIBILITY THAT N_2^+ IS PRODUCED PRIMARILY IN THE *F*-REGION BY SUNLIGHT

We have seen that if the ionization is produced by sunlight, then N_2^+ will recombine in the evening within a few minutes after the appearance of the true shadow in the ionizing wavelength. (With increasing altitude α_{rec} decreases slightly because of the increasing temperature, but n_e increases, so that τ_{rec} remains small.) In this case the screening height is effectively that appropriate to the ionizing radiation. In the morning the N_2^+ emission that is observed will again depend on the height above which N_2 is being ionized. The height of the blue shadow would then have nothing to do with the emitted intensity. To see whether the twilight airglow might arise in this way at very high altitudes, we have computed the overhead equilibrium abundance of N_2^+ for a range of altitudes and for two times after sunset, corresponding to angles $\alpha = 13°$ and $\alpha = 20°$.

4.1. *Ionization equilibrium*

In ionization equilibrium, we have at altitude z,

$$\int_{\nu_i}^{\infty} \pi \mathscr{F}_\nu T_\nu(z) \sigma_\nu(N_2) n(N_2|z)\, d\nu = \alpha_{rec}(z)\, n(N_2^+|z)\, n_e(z), \qquad (5)$$

where n represents number densities, $\sigma_\nu(N_2)$ is the ionization cross-section of N_2, ν_i is the threshold wavelength for N_2 ionization (λ 795), $\pi \mathscr{F}_\nu$ is the solar flux outside the atmosphere in photons cm^{-2} sec^{-1} (unit frequency interval)$^{-1}$, and $T_\nu(z)$ is the *transmission* function or fraction of the incident radiation that penetrates to height z; it is a function of the angle of solar depression as well. In the case of incident line emission, T_ν and $\sigma_\nu(N_2)$ may be removed from the integral, and $\int \pi \mathscr{F}_\nu\, d\nu$ is simply the total line flux in photons cm^{-2} sec^{-1}. Equation (5) is to be solved for $n(N_2^+|z)$; we discuss the adopted data in turn.

For the solar radiation, we consider four principal sources, the ultraviolet continuum, the λ 304 (He II) and λ 584 (He I) line emissions, and soft X-rays. Longward of Ly α (1215 Å) the intensity of the ultraviolet continuum approximates that of a black body at about 4200°K[13]; even if the equivalent temperature in the 100 Å region just shortward of λ 795 is as high as 7000°K, as suggested by de

Jager[14], its contribution will be less than that of the helium lines. For soft X-rays, recent interpretations of rocket observations indicate a total flux of approximately 0·1 erg cm^{-2} sec^{-1} centred about a maximum at 50 Å[15,16,17]. Perhaps 0·02 erg cm^{-2} sec^{-1} is present shortward of 30 Å, but the uncertainty in these results amounts to perhaps an order of magnitude.

For He I λ 584 Rense[18] derives from rocket observations a flux of about 3×10^9 photons cm^{-2} sec^{-1}. For He II λ 304 Rense quotes 15 erg cm^{-2} sec^{-1} (2×10^{11} photons cm^{-2} sec^{-1}); but Hintregger[19] from similar observations derives a flux for the same resonance line which is a factor of 10^3 less. Resolution of this discrepancy is of paramount importance to our problem; in the following we carry both values. The helium lines appear to dominate the solar spectrum between Ly α and the X-ray region[18] and it is unlikely that any important electromagnetic emissions have been neglected.

attained in their experiments, we have assumed a $T^{-1/2}$ dependence[11], which is reasonably consistent with the temperature dependence of α_{rec} measured for $T < 300°K$. Between 300 and 400 km, $\alpha_{rec} = 3 \times 10^{-7}$ cm^3 sec^{-1} is probably a good average.

A model atmosphere was taken from Nicolet[22]; it is based on the assumption of diffusion equilibrium of atmospheric constituents above 110 km and conductive heat transport at high altitudes, fitted to an assumed temperature of 2250°K at 500 km. The model compares well with others based on satellite observations[23, 24], but it should be emphasized that considerable uncertainties still remain.

For the $n_e(z)$ distribution in early twilight, we have adopted data from the extensive tabulations of Thomas *et al.*[25] below the peak of the *F*-region (about 250 km). Above this height we have extrapolated with the satellite data[26, 27] scaled to fit at 250 km. The adopted

Table. 1. High Intensity Solar Radiation and Absorption Coefficients Shortward of 795Å

Radiation	Wavelength (Å)	Flux (Protons cm^{-2} sec^{-1})	$\sigma_a(N_2)$ cm^2	$\sigma_a(O)$ cm^2	h_0 (km)
Ultravoilet continuum ($T = 7000°K$)	650–750	$\sim 2 \times 10^9$	1×10^{-17}	1×10^{-17}	370
He I	584	$\sim 3 \times 10^9$	1×10^{-17}	1×10^{-17}	370
He II	340	$\sim 2 \times 10^{11}$ (Rense)	4×10^{-18}	4×10^{-18} to 8×10^{-18}	300 to 350
		$\sim 2 \times 10^9$ (Hintregger)			
X-rays	50	$\sim 3 \times 10^8$	$\sim 8 \times 10^{-19}$	$\sim 4 \times 10^{-19}$	~ 150
	25	$\sim 6 \times 10^7$	$\sim 1 \times 10^{-18}$	$\sim 3 \times 10^{-19}$	~ 140

Table 1 presents the incident radiation fluxes, along with N_2 ionization cross-sections and O absorption cross-sections from the compilations of Watanabe[20] and Nicolet[21].

Values of α_{rec} are obtained from the measurements of Bialecke and Dougal[12], extrapolated to low pressure according to their linear plot. For temperatures above 300°K, the maximum

O, N_2 and electron distributions are shown in Fig. 2.

4.2. *Transmission function and geometry*

The useful concept of a screening height is equivalent to approximating the transmission function, $T_\nu(z)$, by a step function; i.e. we take $T_\nu(z \geq z_0) = 1$ and $T_\nu(z < z_0) = 0$, where z_0 is

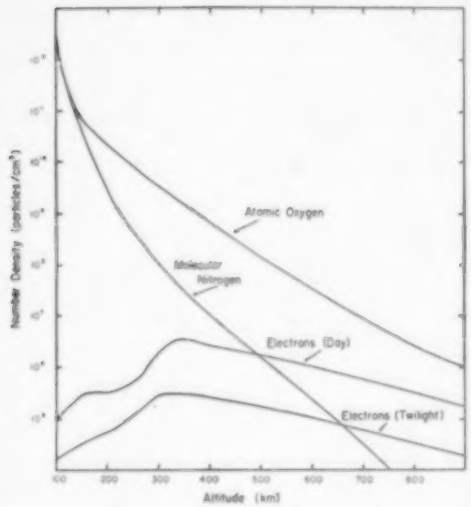


Fig. 2. Adopted atmospheric model.

the height of atmospheric illumination corresponding to a screening height h_0 (cf. Fig. 1). We have used used the concept of a sharp shadow, cast by an atmospheric screening layer, to make preliminary estimates of the ionization produced by various solar radiations. However, in order not to underestimate this ionization production, we have evaluated h_0 as the height corresponding to $T_\nu(z)=1/10$ and listed the results in Table 1.

For a ray that has its closest approach to the earth's surface at altitude h, the transmission function is[28]

$$T_\nu(z)=\exp\left[-\sigma_\nu(\mathrm{O})\,n(h)\,(2\pi aH)^{1/2}\right], \quad (6)$$

provided that z is not approximately equal to h. Here a is the earth's radius and $\sigma_\nu(\mathrm{O})$, $n(h)$ and $H(h)$ (scale height at height $z=h$) all apply to the absorbing substance, which is primarily O for $h>150$ km. We take $n(h)=n(\mathrm{O}|h)$ from Nicolet's model[22] and $H(h)$ from a linear interpolation in Nicolet's model. Setting $T_\nu(z)=1/10$, we obtain the values of h_0 given in Table 1. It is clear that none of the listed radiations can penetrate to the E-layer.

With the screening-height approximation, we would evaluate $n(\mathrm{N}_2^+|z)$ only for $z>z_0$; for a more precise calculation we have computed $T_\nu(z)$ and evaluated the integral in equation (5) for a range of z. In either case, in order to use equation (6), it is necessary to relate the height z to the height of closest approach, h, for a given observation; this is accomplished by the following set of equations (see Fig. 1):

$$\sin\theta=\frac{a}{a+z}\sin\zeta, \quad (7)$$

$$\beta=\theta+\alpha-\zeta, \quad (8)$$

$$\cos\beta=\frac{a+h}{a+z}. \quad (9)$$

For $\alpha=13°$, equations (7) to (9) show that at $z>250$ km, h is approximately equal to z. In this case the line of sight intersects the incident ray near the point of its closest approach to the earth (see Fig. 1) and Hunten's[28] formula for the transmission function must be modified as follows:

$$T_\nu(z)=\exp\left\{-\sigma_\nu(\mathrm{O})\,a\,n(h)\times\int_{-\infty}^{\beta}\exp\left[-a\beta'^2/2H\right]\mathrm{d}\beta'\right\}.$$

The upper limit of integration is here β rather than $+\infty$ as in Hunten's treatment. Separation into two integrals (one from $-\infty$ to 0, the other from 0 to β) and Taylor expansion of the second integrand yield, for $\beta\ll 1$,

$$T_\nu(z)\simeq\exp\left\{-\sigma_\nu(\mathrm{O})\left[(\pi aH/2)^{1/2}\,n(h)+a\beta n(h)\right]\right\} \quad (10)$$

4.3. *Results*

With either equation (6) or equation (10), whichever is more appropriate, and equation (5), the run of N_2^+ number density with altitude can be found. The results are plotted in Fig. 3. Numerical integration of the number density with altitude gives the number of molecules in a square centimeter column; the theoretical and observational values are compared in Table 2. The range of parameters in the λ 304 column is due to the uncertainty in the absorption cross-section of atomic oxygen

Table 2. Observed and Computed λ3914 Intensities (Referred to the Zenith) and Associated Number Densities N_2^+

α for $\zeta=75°$	z_a (km)	Observed*		Computed			
				λ304 ionization ($\pi_\nu \bar{\mathcal{F}}_\nu = 15$ erg cm^{-2} sec^{-1})		X-ray ionization ($\pi_\nu \bar{\mathcal{F}}_\nu = 0{\cdot}1$ erg cm^{-2} sec^{-1})	
		$\mathcal{N}^+(z_a)$ (ions/cm^2 vertical column)	$4\pi\mu_0 \bar{j}$ (rayleighs)	$\mathcal{N}^+(z_a)$	$4\pi\mu_0 \bar{j}$	$\mathcal{N}^+(z_a)$	$4\pi\mu_0 \bar{j}$
13°	100	5×10^9	1000	2×10^9 to 4×10^9	300 to 700	3×10^7	5
20°	200	$1{\cdot}5\times10^9$	30	3×10^8 to 5×10^8	60 to 100	1×10^6	0·2

*Estimated maximum values for quiet magnetic conditions.

at 304 Å[21]. If Rense's λ 304 flux of 15 erg cm^{-2} sec^{-1} is adopted, it is seen that the computed and observed airglow intensities agree reasonably well, in view of the uncertainty of many of the numerical values used. On the other hand, if Hintregger's flux of 0·01 erg cm^{-2} sec^{-1} is correct, we must conclude that solar electromagnetic radiation is incapable of producing the observed atmospheric N_2^+ abundances. Table 2 also shows explicitly the inadequacy of solar X-rays in ionizing molecular nitrogen.

The day airglow may be estimated by a similar procedure, but with the simplification that in equation (5) the transmission function for an overhead sun is now given by $T_\nu(z) = \exp[-\sigma_\nu(O)\, n(z)\, H(z)]$. Fig. 3 gives the resulting run of $n(N_2^+|z)$. The total N_2^+ abundance, again if the λ304 flux is 15 erg cm^{-2} sec^{-1}, is $3{\cdot}6\times10^{11}$ molecules cm^{-2}, corresponding to a zenith emission rate of $4\pi\mu_0 \bar{j} = 70$ kR. In this case, the λ3914 dayglow would be among the most intense in the visible region of the spectrum. On the other hand, if the λ304 flux is 0·01 erg cm^{-2} sec^{-1}, $4\pi\mu_0 = 50$ R. This difference should be easily detectable, and it appears that a rocket observation of the λ3914 dayglow could distinguish between the two alternative λ304 fluxes which have been proposed. The λ3914 dayglow produced by X-ray ionization and resonance scattering has a zenith emission rate of about 700 R.

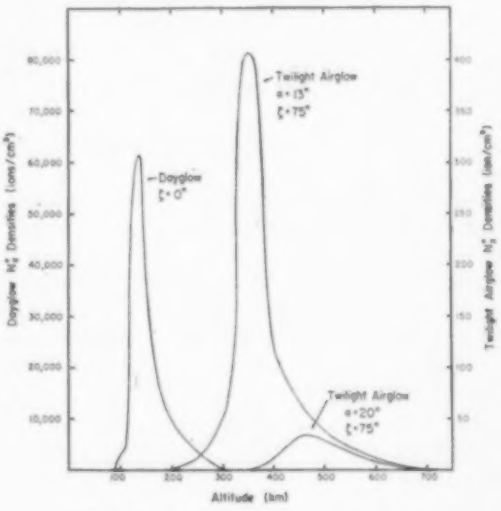


Fig. 3. Calculated N_2^+ abundances.

5. THE POSSIBILITY THAT N_2^+ IS PRODUCED PRIMARILY IN THE *E*-REGION BY PARTICLE BOMBARDMENT

In the case that the lower λ 304 flux is the true one, we have seen that solar electromagnetic radiation cannot account for even the quiet-day N_2^+ twilight emission. And, as Bates[5] pointed out, the enhancement during magnetically active periods requires an additional source of ionization. Particle bombardment seems the most likely alternative.

We shall suppose, for lack of anything more

definite, that the particle flux is constant during twilight. The screening height is then that applicable to λ 3914 and the true emission height is in the *E*-region. From Table 2, we conclude that the integrated abundance of N_2^+ above a true height of about 115 km is normaly $N^+ < 5 \times 10^9$ ions cm^{-2}. The intensity does seem to increase when the shadow drops below this altitude, so for the entire atmosphere this upper limit may well be approached even under quiet conditions. With $n_e \approx 1{\cdot}5 \times 10^4$ electrons cm^{-3} and $\alpha_{rec} \approx 7 \times 10^{-7}$ cm^3 sec^{-1} for the twilight *E*-region, the rate of N_2 ionization in the entire column is $\alpha_{rec}\, n_e \approx 5 \times 10^7$ ions cm^{-2} sec^{-1}. If the ionization is produced by electrons, 30 to 35 eV will be lost in producing each ion pair[29]. Allowing for ionization of O and O_2 in addition to N_2, we find that a total energy flux for the incident electrons of about 3×10^9 eV cm^{-2} sec^{-1} is required. Electrons penetrating to *E*-layer depths (110 km) will have an initial energy of about 3 keV[9]. Hence the required electron flux is about 10^6 electrons cm^{-2} sec^{-1}. For comparison, Van Allen[30], in a summary of rocket results, concludes that the electron flux in the auroral zone and in the energy range 10 to 100 keV is approximately $10^6 - 10^8$ electrons cm^{-2} sec^{-1}.

If the nitrogen ionization is produced by protons, an initial energy of about 30 keV is required for the particles to reach the depth of the *E*-region[9]. Protons also lose about 33 eV per ion pair[29]. We then find that the necessary proton flux is only about 10^5 protons cm^{-2} sec^{-1}. In the aurora, proton fluxes exceeding 10^7 protons cm^{-2} sec^{-1} are not unusual according to the Hα intensities[31, 32].

If one ionization in fifty also produces a quantum of λ 3914, we would expect a zenith emission rate of $4\pi\mu \mathcal{J}_{3914} \approx 1$ R. There may be a weak (perhaps a few rayleighs) λ 3914 emission in the nightglow; a component of only 1 R would not be detected over the other radiations. However, during disturbed magnetic conditions the N_2^+ abundance is evidently many times greater than the value adopted here. It thus seems not impossible that λ 3914 may appear at times in the nightglow as a result of particle bombardment, perhaps constituting a weak permanent aurora emanating primarily from the *E*-region. Possibly these bombarding particles are the leakage of the Van Allen radiation from magnetically trapped orbits.

REFERENCES

1. V. M. Slipher, *Mon. Not. R. Astr. Soc.* **93,** 657 (1933).
2. C. T. Elvey, *Re. Mod. Phys.* **14,** 140 (1942).
3. M. Dufay, *Ann. Phys.* **8,** 813 (1953).
4. J. and M. Dufay, *C. R. Acad. Sci., Paris,* **224,** 1834 (1947).
5. D. R. Bates, *Proc. Roy. Soc.* A**196,** 562 (1949).
6. O. R. Wulf and L. S. Deming, *Terr. Magn. Atmos. Elect.* **43,** 283 (1938).
7. J. W. Chamberlain, *Astrophys. J.* **127,** 54 (1958).
8. H. Shull, *Astrophys. J.* **112,** 352 (1950).
9. D. R. Bates, *The Earth as a Planet* (edited by G. P. Kuiper), Chap. 12. University of Chicago Press (1954).
10. P. Swings and M. Nicolet, *Astrophys. J.* **109,** 327 (1949).
11. D. R. Bates, *Phys. Rev.* **78,** 492 (1950).
12. E. P. Bialecke and A. A. Dougal, *J. Geophys. Res.* **63,** 539 (1958).
13. R. Tousey, *The Sun* (edited by G. P. Kuiper), Chap. 9.9. University of Chicago Press (1953).
14. C. de Jager, *Ann. Geophys.* **11,** 330 (1955).
15. G. Elwert, *Z. Naturf.* **9**a, 637 (1954).
16. H. Friedman, *Ann. Geophys.* **11,** 174 (1955).
17. E. T. Byram, T. A. Chubb and H. Friedman, *J. Geophys. Res.* **61,** 251 (1956).
18. W. A. Rense, private communication (1958).
19. H. Hintregger, private communication (1959).
20. K. Watanabe, *Advances in Geophys.* **5,** 154 (1958).
21. M. Nicolet, *Physics and Medicine of the Upper Atmosphere* (edited by C. S. White and O. O. Benson), Chap. 12. University of New Mexico Press, Albuquerque (1952).
22. M. Nicolet, *Ann. Geophys.* **15,** 1 (1959).
23. T. E. Sterne, B. M. Folkart and G. F. Schilling, *Spec. Rep. Smithsonian Astrophys Obs., Camb.* No. 7, IGY Project No. 31.10 (1957).
24. I. Harris and R. Jastrow, *Planet. Space Sci.* **1,** 20 (1959).
25. J. O. Thomas, J. Haselgrove and A. R. Robbins, *Tables of Ionospheric Electron Density, Slough (England).* Cavendish Laboratory, Cambridge (1957).

26. J. W. TOWNSEND, *Composition of the Upper Atmosphere*, paper presented at the Second International Symposium on the Physics and Medicine of the Atmosphere and Space, San Antonio, November 1958. To be published.
27. J. C. SEDDON and J. E. JACKSON, *Ann. Geophys.* In press.
28. D. M. HUNTEN, *J. Atmos. Terr. Phys.* **5**, 54 (1954).
29. J. M. VALENTINE and S. C. CURRAN, *Reports on Progress in Physics* Vol. 21, p. 1. Physical Society, London (1958).
30. J. A. VAN ALLEN, *Proc. Nat. Acad. Sci., Wash.* **43**, 57 (1957).
31. J. W. CHAMBERLAIN, *Astrophys. J.* **120**, 360 (1954).
32. J. W. CHAMBERLAIN, *Astrophys. J.* **126**, 245 (1957).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 165–173. Printed in Great Britain.

PROPERTIES OF THE UPPER ATMOSPHERE AND THEIR RELATION TO THE RADIATION BELTS OF THE EARTH

S. F. SINGER*

Physics Department, University of Maryland, College Park, Maryland

(*Received* 29 *June* 1959)

Abstract—A new theory has been developed which gives the distribution of density with altitude for a planetary exosphere in the absence of thermodynamic equilibrium. The results differ considerably from those calculated on the basis of a hydrostatic equation.

To apply this theory to the earth's exosphere, we first fix the temperature of the base of the exosphere, which is located at 530 km, by using satellite drag data between 500 and 1000 km. A temperature of 1500° gives a reasonable fit to the rather sparse data and allows us to calculate the atomic oxygen distribution with altitude.

From the cosmic ray neutron albedo theory of the inner (cosmic ray) radiation belt, we can set an *upper limit* of about 10^6 hydrogen atoms per cm^3 at an altitude of 1000 km, but their concentration could be less. However, the total density including the ionized component must be of the order of 10^{-19} g/cm^3 at 1000 km.

The existence of a minimum between the two observed radiation belts, the so-called "slot", is accounted for in terms of the breakdown of the adiabaticity of the magnetic moment of trapped particles. This same theory also allows one to make statements concerning the mean energy of the trapped protons in the inner radiation belt. It should become progressively softer with increasing altitude.

The outer (auroral) radiation belt, which is centered at 3 earth radii, has a lifetime which is controlled primarily by the presence of a scattering atmosphere. From a theory which takes into account both energy loss and scattering, one can calculate the intensity contours for the quasi-equilibrium state. Special techniques give the transient case following the injection of particles into this region.

The initial injection of solar corpuscular radiation consists mainly of protons and electrons with velocities of 2×10^8 cm/sec. These 20 keV protons form a third and rather transient ($\sim$ 1 day) radiation belt at 4–8 earth radii; they account for most of the effects observed during magnetic storms. Their initial diamagnetic effects produce the "reverse sudden commencement" (*SC**). Their increase of upper atmosphere ionization accounts for the sudden commencement currents while their drift in longitude accounts well for the main phase of magnetic storms. In comparing various removal processes for the injected protons, we find that charge exchange is about 100 times more effective than atmospheric scattering. Using our theoretical distribution for the hydrogen exosphere, we can therefore calculate the complete development of a magnetic storm as a function of time, including the distribution around the earth of trapped protons as well as their magnetic effects.

* This research was supported in part by the United States Air Force under Contract AF 49(638)-530 monitored by the Air Force Office of Scientific Research of the Air Research and Development Command and by the Electronics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract No AF 19(604)-5575.

VARIATION OF AIR DENSITY IN THE EXPOSURE

In any planetary atmosphere one reaches a level of altitude at which the density of the atmosphere is so low that the mean free path becomes of the order of the scale height. To a good approximation one may assume that above this level the mean free path in fact becomes infinite so that atoms which travel in this region, named the *exosphere*, would experience no collisions with other atoms. It can be easily seen then, that atoms which issue from the base of the exosphere will describe ballistic orbits, i.e. portions of ellipses which will take them out to high altitudes before they again re-enter the atmosphere at a different point. Some of the atoms will in fact have velocities high enough, greater than the escape velocity, which will take them beyond the gravitational field of the planet.

For the planet earth, we find the critical level h_c to be at 530 km. At h_c we must have a concentration n_c so that $n_c = (4\pi a^2 H_c)^{-1}$; with $a \sim 1{\cdot}5 \times 10^{-8}$ cm we find $n_c \sim 4 \times 10^7$ cm^{-3} = n (0); i.e. we have practically only oxygen atoms. Below about 350 km, N_2 becomes an important constituent; at 220–230 km it predominates.

We may therefore take the 530-km level as the base of the exosphere. For more accurate considerations, one must take into account that at that level only the mean free path in the horizontal direction is equal to the scale height. In reality, the mean free path is not isotropic and is therefore a function of the angle which the velocity vector of the particle makes with the local vertical.

It is of great interest for many problems to know the density of neutral particles in the exosphere. For example, in order to calculate the drag on a satellite or other space vehicle, or to calculate the lifetime of trapped radiation belt particles, we need to know both the neutral and ionized particle density. While the ionized particle density is controlled by the earth's magnetic field and other factors, the concentration of neutral particles is controlled entirely by the conditions existing at the base of the exosphere and the gravitational field of the earth, which of course is well known. Starting with these two ingredients one can calculate the exospheric densities provided one has a theory which gives the density distribution.

A thermodynamic approach to this problem has usually been attempted.[1] It has been assumed that the gas in the exosphere is isothermal and in fact has the same temperature as the gas at the base of the exosphere.* Using then the idea of an "isothermal" exosphere, one can calculate the density distribution very simply from thermodynamic considerations. It is given by the usual hydrostatic equation

$$n(r) = n_0 \exp\left[\frac{GMm}{kT}\left(\frac{1}{r} - \frac{1}{r_0}\right)\right] \qquad (1)$$

Further reflection shows that this theory cannot be correct. In the first place, an inspection of equation (1) shows that the density at infinity does not drop to zero but remains finite. Secondly, if we consider for a moment the situation at very large distances from the earth, then we should find there practically only particles which at the base of the exosphere have had escape velocity or greater. These particles of course are destined to escape and will form the escape flux at *any* level. At very large distances from the earth these particles will be streaming outward with nearly radial velocity. It is clear therefore, that the velocity distribution of the particles is not isotropic and one may therefore raise the question concerning the significance of the concept of a temperature. In fact, in the absence of collisions between particles, which is of course assumed in the exosphere, the assumption of thermodynamic equilibrium which underlies the hydrostatic equation is very much in doubt.

We have decided to forego the assumption of thermodynamic equilibrium which cannot be correct in any case, and calculate the density distribution by a different method, namely by

* Intuitively one might think that the temperature of the gas decreases as the velocity decreases when particles rise to higher altitudes. However, it is also intuitively clear that particles of very low velocity do not reach to very high altitudes and that therefore the mean square velocity should not vary too greatly with altitude.

considering the statistical distribution of particle orbits issuing from the base of the exosphere and integrating their contribution to the density of each level.

escaping fraction of the atmospheric gas is all that contributes to the density. It is clear therefore that at large distances from a planet, the density must fall as $1/r^2$.

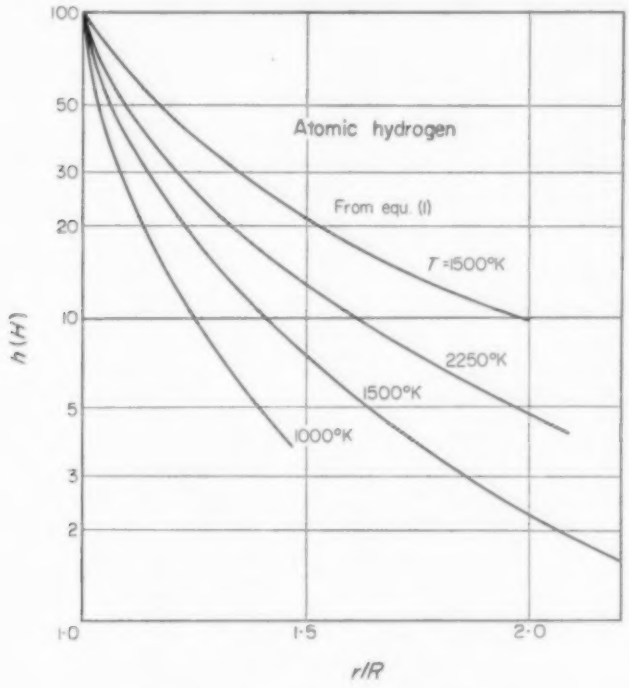


Fig. 1. Relative concentration of hydrogen atoms for different temperatures calculated according to Reference 2 as a function of distance from center of the earth. For comparison we show the distribution according to equation (1) for T=1500°K.

The details of the calculation are given in Ref. (2). However, we may compare the density distribution calculated from the new theory with that given by equation (1) (see Fig. 1). The difference is considerable; the density immediately falls off much faster above the base of the exosphere, it is consistently less at all levels. For example, at a distance of 2 earth radii, the logarithmic densities differ by 0·6, at 5 earth radii they differ by 2·37, or by more than two orders of magnitude.

At very large distances from the earth the

At lower altitudes however, the contribution of the ballistic orbits which again return to the earth is of much greater significance and therefore the density distribution is in fact quite complicated. We have attempted to solve this problem by methods involving series expansion of integrals and by numerical integrations.

Temperature

We are referring strictly to the temperature at the base of the exosphere. One way of finding it is to study experimentally the distri-

bution of density as a function of altitude. At altitudes below 1000 km it is reasonably certain that all of the contribution to the density comes from atomic oxygen. We can therefore use satellite drag data, of which there is very little, to calculate the temperature in this region. We do this by normalizing the theoretical distribution to the point determined by Harris and Jastrow at 720 km.[3] We find little difference between 1250° and 1500°K (see Fig. 2), but

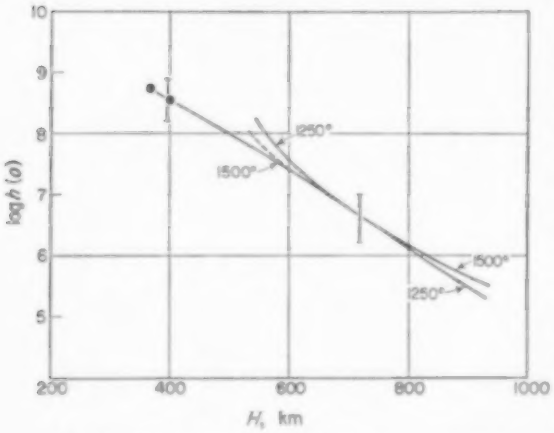


Fig. 2. The concentration of atomic oxygen as a function of altitude in the inner exosphere derived from satellite data. It can be seen that both 1250° and 1500° provide good fits to the two data points below the base of the exosphere. More data points with higher precision would seem definitely to be required.

1500°K fits in much better with the general trend of densities at lower altitudes and also of course with the escape rate of helium, as has been pointed out by Spitzer[3]. It is essential to extend the satellite drag data in the region 530–1000 km.

HYDROGEN CONCENTRATION

Integral method of normalization

One way of finding the hydrogen density as a function of altitude, knowing the theoretical distribution calculated earlier, is to determine the total amount of hydrogen *above* a certain level. This approach has been taken by F. S. Johnson[4] using the Lyman-α contours obtained by Friedman and collaborators.[5] He has assumed a thin atmosphere and then shown that the data are consistent with such a thin atmosphere. However, if the atmosphere is really thin, then the hydrogen close in to the earth extending out to about 0·4 of an earth radius is not illuminated and does not contribute. Conceivably, therefore, a thick atmosphere could explain the data equally well. In fact, we cannot as yet definitely exclude interplanetary hydrogen as a source of the night time Lyman-α radiation, coupled with an optically thick hydrogen exosphere. The experimental data simply are not good enough to allow a firm decision. One needs to measure line widths down to 0·01 Å in order to determine whether the light observed is scattered from terrestrial hydrogen or interplanetary hydrogen.

Differential method using radiation belts

Another way to normalize the theoretical hydrogen distribution is by a direct measurement of hydrogen density at one or preferably more than one level. The radiation belts give us a means of at least estimating the hydrogen densities and we will apply theoretical discussions of the radiation belt intensities to this problem in order to learn as much as possible about the hydrogen densities above 1000 km.

(i) *Inner (cosmic ray) belt.* By comparing the observed radiation belt flux at an altitude of 1900 km with the flux expected on the basis of the cosmic-ray albedo neutron theory,[6] we can derive an *upper limit* to the hydrogen density which is 9×10^5 atoms/cm^3. It should be noted that this is an upper limit. If other processes besides atmospheric collisions shorten the lifetime of the trapped particles then the density could be less.* Also, if the flux of albedo neutrons is less than that calculated, which is possible, then again the density of the atmosphere at that level could be reduced by a factor of 2. From these considerations we could therefore stretch the

* In fact, the gradual loss of adiabaticity of the particles' magnetic moment reduces this density by a factor of 3 (see discussion in next section).

present calculation to allow a *minimum* density of the order of 10^5 neutral atoms/cm^3 *or* $2{\cdot}5\times10^4$ ionized hydrogen atoms per cm^3*. However, whistler data indicate the existence of an ionized atmosphere of about 2×10^4/cm^3 at this level so that our method loses sensitivity for the determination of the neutral hydrogen density if we go to the lower limit.

A theory for the "radiation belt slot"

In view of experimental data[7] which show a maximum for the inner radiation belt at an altitude of about half an earth radius it may be questioned whether in fact its intensity is controlled by atmospheric density. We have assumed earlier, and we will show that this assumption is essentially correct, that the intensity is controlled by atmospheric density at low altitudes[8] which is clearly shown by its rapid rise. However, as we pointed out, the adiabatic invariance of the particle's magnetic moment must break down at a certain point. What this point is cannot be decided from *a priori* considerations. However, the work of Northrop indicates that the breakdown is quite sharp. We now associate a maximum momentum and therefore a β_{max} with each altitude. These considerations lead to a breakdown law given by[6]

$$\beta_{max}=0{\cdot}7r^{-2} \qquad (2)$$

By applying this law in addition to the calculated hydrogen density d we find for the omnidirectional integral radiation belt flux at the equator[6]

$$I=8{\cdot}6\times10^{-14}\,(\eta/\pi)\cos\alpha_c[\beta_{max}{}^{1{\cdot}4}-\beta_{min}{}^{1{\cdot}4}]$$

$$\text{d}^{-1}/\text{cm}^2\text{-sec} \qquad (3)$$

where η is the mean injection factor $\simeq0{\cdot}4$; all other symbols are as before.[6] Without any further assumptions the flux shows a maximum at $1{\cdot}5R_E$, which corresponds closely to the observations (see Fig. 3).

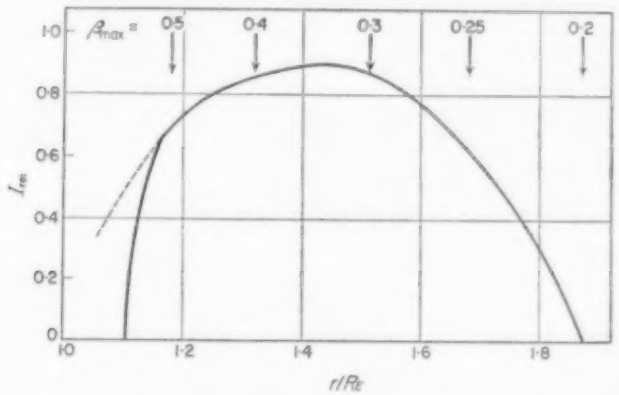


Fig. 3. Omnidirectional integral intensity (relative) of cosmic ray radiation belt near equator. β max varies as indicated ; β min is taken as 0·2 (20 MeV protons). The dashed line shows the intensity variation expected if only hydrogen existed in the exosphere. The presence of oxygen, however, causes the steep initial rise shown. The transition from O to H, we calculate, should take place at about 1000 km. However, it will be interesting to obtain an experimental verification of this.

* It can be shown that an ion pair is 4 times as effective as a neutral atom against nonrelativistic protons.

It might be mentioned parenthetically that if this theory for the "slot" is correct, then the Argus electrons which extend out as far as 2 earth radii in the equatorial plane are well within the adiabatic limit and their lifetime is therefore controlled by atmospheric density.

The theory which gives the breakdown of the adiabatic condition as a function of altitude can be easily verified. It should lead to a progressive diminution of β_{max} in accordance with equation (2) and, as seen from Fig. 3, this would lead to a progressive softening of the radiation as one moves out to higher altitudes. For example, at 1·5 earth radii there should be no protons present with energies greater than about 50 MeV; and at 1·8 earth radii no protons with energy greater than about 25 MeV. This implies also that the amount of shielding required at various altitudes for manned space stations diminishes with increasing altitude.

(ii) *The auroral belt.* As shown by the considerations on adiabatic invariance developed in the preceding paragraph, the particles in the auroral belt which we believe consist of 50 keV electrons (ranging perhaps down to 10 and up to 500 keV) will satisfy the adiabatic condition quite well and in the absence of magnetic field perturbations their lifetime would be controlled strictly by the presence of atmospheric ions, electrons and neutral atoms. A theory for the lifetime of trapped particles has been developed[9] using the Fokker–Planck equation to take account both of the scattering and of the energy loss of particles. From this theory we can derive the quasi-equilibrium distribution of intensity in the belt (see Fig. 4).[9]

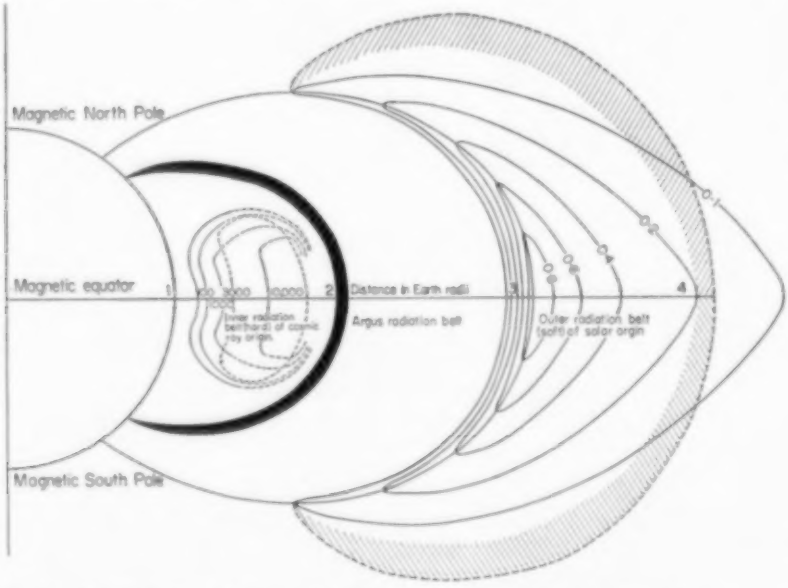


Fig. 4. The radiation belts around the earth. The experimental regions for a maximum of the inner and outer radiation belt are shown according to Van Allen. For the inner radiation belt we show the calculated contours from the neutron albedo theory (Ref. 6). The Argus radiation belt is shown at two earth radii. For the outer radiation belt we show contours calculated from considerations of lifetimes for trapped electrons according to Ref. (9). The agreement is fair between these and contours which have been estimated from observations so far.

Clearly, if the particles are injected with a peculiar type of angular distribution, then a transient decay will occur first which rapidly disappears and finally goes into a quasi-equilibrium exponential decay. For example, the injection of the Argus electrons took place in a very peculiar manner leading to a peculiar distribution of pitch angles. It is therefore thought that the *observed* decay of the Argus electrons was the transient decay rather than the quasi-equilibrium decay which we have calculated.[10]

It is unfortunately not possible to calculate the density of the atmosphere from the observed auroral belt intensities. This is so because we cannot calculate *a priori* the injection rate into the auroral belt, the rate of supply of particles. (In the case of the cosmic ray belt, this is possible from the cosmic ray neutron albedo theory.) However, it is possible to obtain these densities by artificially injected electrons.[11] One would prefer electrons from an accelerator, injected with a well known energy and injected either isotropically or in a known direction.

(iii) *The magnetic storm belt.* We have evaluated the magnetic effects produced by the drift of trapped particles in both inner and outer radiation belts and find that they are not adequate to explain magnetic storm effects. We therefore must go back to our original hypothesis[12] that protons with velocity of 2×10^8 cm/sec, having energies of approximately 20 keV, are captured in the earth's magnetic field after about one day following a solar flare. Injection takes place as the solar plasma cloud interacts with the weak geomagnetic field, and "kinks" it as instabilities are set up. The "kinks" can then scatter particles in the trapping region, which—as Störmer has shown—is normally inaccessible.[13]

Since this injection is reasonably well-defined and presumably lasts of the order of a few hours, we can attempt to calculate the density of the atmosphere in the region occupied by these protons from their observed lifetime. The data on the main phase of magnetic storms give a mean life of about one day.

Again we must investigate what factors determine the lifetime of the trapped 20 keV protons. As seen from equation (2) the adiabatic invariance is satisfied out to about 8·5 earth radii. The effects of the atmosphere include collision losses and scattering, as well as charge exchange. To determine which of these effects is most important we compare the experimentally determined charge exchange cross section[14] of 7×10^{-16} cm^2 with a calculatated collision and scattering coefficient for 20 keV protons of 2×10^{-9} cm^3/sec.[9] We see therefore that charge exchange is about 100 times as effective as scattering in removing the low energy trapped protons. However, we must assume that the ratio of ionized to neutral hydrogen is not greater than 100.

The decay time of magnetic storms thus determines the neutral hydrogen density at about 100 per cm^3. Its location corresponds to that of the very transient magnetic storm belt. We would place it about 5 to 8 earth radii in order to account for the observed value of the magnetic storm decreases during the main phase. We have earlier calculated the drift velocity of these particles and shown that the drift in longitude of the protons trapped in this region could produce a current of the required magnitude.[12] The diamagnetic effects of the protons' gyration will decrease the effect by about one-third but cannot cancel it completely.

Geophysical effects

We associate this transient belt with magnetic storms as well as with aurora. In our view, the outer radiation belt found in the Pioneer and Mechta rockets is not really directly responsible for most of the aurora since otherwise it would not show such a long lifetime. After the bulk of the protons are removed by charge exchange, the remaining protons and a fraction of the electrons are accelerated locally, i.e. within the earth's magnetic field, to high enough energies to produce the observed aurorae. The acceleration mechanism has not been elucidated in detail. We have discussed a type of Fermi-acceleration in which the trapped particles interact with magnetohydrodynamic waves, which are themselves loosely trapped in the earth's

magnetic field, by being reflected from the upper ionosphere at an altitude of several hundred kilometers.[15] Then from purely thermodynamic considerations one can show that the trapped particles are accelerated as they tend towards equipartition; in other words, as their kinetic energy tends towards the kinetic energy contained in the magnetohydrodynamic waves. However, as they are accelerated their pitch angle is also changed; hence many of the particles are lost from the trapping region and will precipitate into the auroral zone.[15]

It is interesting to study the geophysical effects of the transient radiation belt made up of the 20 keV protons. Initially the pitch angle distribution will be flat. In other words, many protons will be able to get down close to the earth (see Fig. 5). These will not drift but their diamagnetic effects will instead push the lines of force apart or weaken the field in the auroral zone. In our view, this feature is responsible for the reverse sudden commencement, the so-called *SC**, which precedes the sudden commencement of magnetic storms, and is

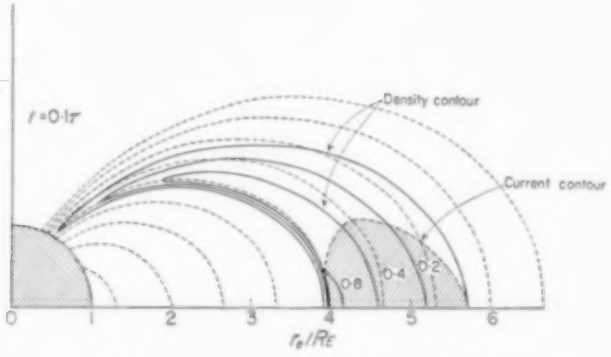


Fig. 5. The concentration of trapped solar protons around the earth at the beginning of a magnetic storm. Note the large number of protons which can get close to the earth where their diamagnetic effects will be important and produce the reverse sudden commencement.

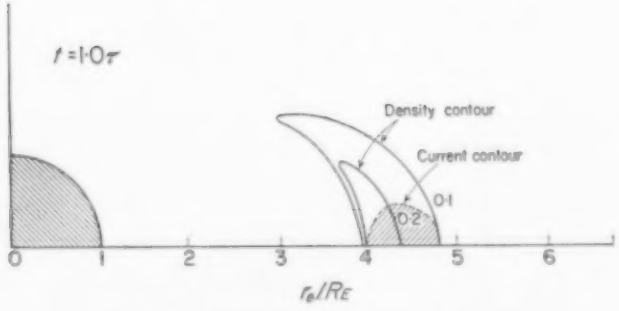


Fig. 6. The distribution of trapped solar protons around the earth during the main phase of the magnetic storm. Here the azimuthal drift is the important quality which determines the magnetic effects observed at sea level.

generally observed in the auroral zone.[12] The sudden commencement itself may be due to the incidence of such protons into the ionosphere, there causing additional ionization which increases conductivity and allows existing dynamo e.m.f.'s to drive a larger current through the ionosphere. Since we know now the atmospheric neutral hydrogen distribution as a function of altitude, we can fairly directly calculate the spatial distribution of these trapped protons as a function of time. We can also calculate the changing magnetic effects observed at sea level as a function of time by following the decay of the whole cloud (see Fig. 5 and 6).[16]

I wish to thank R. C. Wentworth for useful discussions and assistance in some of the calculations.

REFFERENCES

1. L. Spitzer, *The Terrestrial Atmosphere Above 300 km in The Atmospheres of the Earth and Planets* (edited by G. P. Kuiper) pp. 211–247. Univ. of Chicago Press (1952).
2. E. J. Opik and S. F. Singer, *Physics of Fluids*, **2**, 653 (1959).
3. I. Harris and R. Jastrow, *Science*, **128**, 420 (1958).
4. F. S. Johnson, *The Structure of the Outer Atmosphere Including the Ion Distribution Above the F-2 Maximum*. to be published.
5. J. E. Kupperian, Jr., E. T. Byram, T. A. Chubb and H. Friedman, *Planet. Space Sci.* **1**, 3 (1959).
6. S. F. Singer, *Phys. Rev. Letters*, **1**, 181 (1958).
7. J. A. van Allen and L. A. Frank, *Nature, Lond.* 430 (1959).
8. S. F. Singer, *Phys. Rev. Letters*, **3**, 188 (1959).
9. R. C. Wentworth, W. M. MacDonald and S. F. Singer, *Physics of Fluids*, **2**, 499 (1959), *Bull. Amer. Phys. Soc.* **4**, 7 (1959).
10. R. C. Wentworth, private communication.
11. S. F. Singer, *J. Brit. Interplan. Soc.* **16**, 558 (1958).
12. S. F. Singer, *Trans. Amer. Geophys. Un.* **38**, 175 (1957).
13. S. F. Singer, *Bull. Amer. Phys. Soc.* **1**, 229 (1956).
14. W. L. Fite, R. T. Brackmann and W. R. Snow, *Phys. Rev.* **112**, 1161 (1959).
15. S. F. Singer, *Ann. Geophys.* **14**, 433 (1958); *Bull. Amer. Phys. Soc.* **3**, 40 (1958).
16. S. F. Singer and R. C. Wentworth, *J. Geophys. Res.*, to be published.

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 174–186. Printed in Great Britain.

ARTIFICIAL ELECTRON CLOUDS—V

MORNING TWILIGHT STUDY, RELEASE OF CESIUM AND SODIUM AT 128 AND 116 km

F. F. MARMO,* J. PRESSMAN,* E. R. MANRING* and L. ASCHENBRAND*
Geophysics Corporation of America, Boston 15, Massachusetts

(*Received* 20 *July* 1959)

Abstract—This paper, one of a continuing series, reports the results and analysis of two rocket experiments designed to study the simultaneous optical and radio-radar behavior of artificial electron clouds. The optical properties were obtained by "doping" the cloud with a small percentage of sodium in one experiment and using a complete payload of sodium in another. These experiments were performed with the Nike-Cajun under morning twilight conditions on 21 and 22 May 1958 respectively.

Analyses are made of radar slant range behavior with time, optical growth rate and variation of critical frequency with time. In the last, a more advanced analysis is used which allows for the differential diffusion of charged and neutral particles. The results from these analyses are compared to the estimated atmospheric properties and to thermo-chemical release properties. These are also cross-checked among themselves.

This study indicates the utility of combined radio-radar and optical observations. Values are derived for the following parameters: the neutral and ambipolar diffusion coefficients, chemical yield, thermal ionization efficiency, wind velocity and wind shear.

1. INTRODUCTION

This paper is one of a continuing series devoted to the elucidation of the physics of artificial electron clouds. Paper I[1] of this series presented a concise summary of the data obtained in eight rocket experiments in which artificial electron clouds were produced; the initial paper also listed some suggested scientific uses of these clouds. Paper II[2] discussed the general considerations in the overall design and interpretation pertinent to a systematic study of this nature. Paper III[3] presented an analysis of the electron cloud created by the daytime release of atomic potassium at 121 km by means of an Aerobee rocket. Paper IV[4] reported on the electron cloud created by the 101 km nighttime release of atomic cesium from a Nike-Cajun rocket substantiating the thesis that thermal production of electrons was an active process in experiments of this nature.

It has become apparent that a knowledge of the dynamics of the initial expansion and subsequent diffusion of the released atomic species was essential to the understanding of the physics of electron cloud radar returns. It was decided to render the clouds visible by "doping" the chemical payload of cesium compound with a trace of sodium in one case, and by flying a complete sodium payload in a second experiment. The technique was to adopt morning twilight releases since it was essential that the cloud be solar illuminated for an extended period under low intensity sky backgrounds. Previous experiments have shown that the solar illuminated sodium cloud would be easily visible.[5] On the basis of

* Formerly with Geophysics Research Directorate, Air Force Cambridge Research Center, L. G. Hanscom Field, Bedford, Massachusetts.

This work has been partially supported under contracts: U.S. Army Signal Corps DA36-039-SC-78971 and U.S. Army Signal Corps DA36-039-SC-78925.

Rocket experiments were performed when authors were employed at GRD, AFCRC.

similar chemistry and negligible reduced mass effects it was assumed that the cesium atoms would be distributed in a manner proportionate to the sodium tracer.

It was felt that the reduced radio-radar data determinations of thermal-ionization efficiency, chemical yield, diffusion coefficient, etc., could be cross-checked with optical data to yield integrated, interlocked values for the basic processes and parameters.

In a previous paper[3] for the analysis of *C*-3 radar data, a simplified electron cloud model was adopted which ignored the possible effects of differential diffusion. Here, a more sophisticated model is developed which considers the role of the differential diffusion of ion-electron and neutral species on the cloud dynamics.

Additionally in Paper III,[3] because of the logistic requirements of an extended program of rocket experimentation, there was explored the possibility of using rockets of smaller payload, namely, Nike-Cajun. The analysis indicated that the Nike-Cajun was adequate for the purpose of this study. The initial experiment reported here checked this hypothesis.

It should be commented that in this paper as in Paper III, the dominant density decay mechanism is considered to be diffusion and the other loss mechanisms such as recombination, mutual neutralization, etc., are not considered to be significant.

2. EXPERIMENTAL

2.1. *Experimental considerations*

On the basis of previous experimentation, release altitudes were chosen to insure that the only operable electron decay process would be diffusion. These altitudes were found to be approximately 115 and 125 km for sodium and cesium, respectively (see Fig. 6, Paper III[3]).

During twilight periods the ozone layer will cast a shadow in ultra-violet analogous to the earth's optical shadow and somewhat higher. Calculations indicate that for a 50 km ozone shadow[6] releases at these altitudes should be performed about 30 min before local sunrise to assure that ionizing solar radiation reaches the cloud.

The utilization of sodium as an optical "doping" agent is particularly effective. First, its optical cross-section is so high that only about 10^{11} atoms/cm^2 column is required to render a cloud optically dense. Secondly, for alkali metal studies it can be safely assumed that the chemistry is similar to a high degree for all species. For example, it may be assumed that the ratio of cesium to sodium atoms existing in the payload is maintained in the dispersed cloud. Furthermore, since the photoionization probability of sodium is less than 2 per cent that of cesium, no detectable effect due to photoionization would be expected, *a fortiori*, since only a small percentage of sodium was used. Similarly, since the thermal-ionization efficiencies lie in the same direction, no substantial sodium thermal-ionization is expected. These factors serve to justify the use of sodium as an effective optical "doping" agent.

The release of pure sodium of 22 May 1958 had several purposes. The optically observed growth could be used to study cloud dynamics under optimum conditions; also in view of a previous release of potassium the three alkali metals could be compared as to efficiency. Furthermore, predictions made in an earlier paper[3] indicated that the probability of generating detectable ion clouds with a sodium release from a Nike-Cajun rocket was low, and it was desirable to check this fact experimentally.

2.2. *Payload and chemistry*

The basic engineering of the metallic vaporizers employed in these experiments has been previously described in a paper[4] dealing with the release of atomic cesium at an altitude of 101 km. The background and chemistry has also been discussed previously.[5, 4] For these experiments the payload and geometry were modified in order to conform with the capabilities of the Nike-Cajun rocket nose cone. In particular, the empty canisters (400 in^3 volume) weighed about 8·6 kg. The chemicals used in the first experiment reported here consisted of 8·9 kg of cesium nitrate (45·7 gram moles), 4·7 kg of aluminum granules and 0·1 kg of sodium nitrate (added as a "doping" agent). The

$CsNO_3$/Al ratio was 1·9. This payload was launched on a Nike-Cajun rocket on 21 May 1958, at Holloman AFB, New Mexico, at 0431:59 MST. At 118 sec after launch and at 128 km, the chemical constituents were successfully released.

The second experiment, the release of atomic sodium, was performed on 22 May 1958. A Nike-Cajun rocket released the chemical payload 108 sec after launch time (0429:59 MST) at an altitude of 116 km. The loaded vaporizer in this case contained 5·05 kg (61 gram moles) of sodium nitrate and 5·85 kg of aluminum granules.

3. EXPERIMENTAL PROBES AND RESULTS

3.1. *Optical probes*

Details concerning some of the optical instrumentation employed for optical probes have been provided in previous publications of this series.[1, 4] These include: (*a*) the Astro Ballistic camera sites on the HAFB-WSPG range;[1, 4] (*b*) the Baker–Nunn satellite tracking camera[4] at Organ Pass, New Mexico; this camera was fitted with a yellow filter and utilized 4 sec exposures for these two experiments (*c*) the Beattie–Coleman Varitron 35 mm recording camera[4] which was equipped with a narrow band 5890 Å interference filter. In addition to the above, the following probes were utilized and have not been previously described.

A Bendix light amplifying system called a Lumicon was employed. This is an electronic instrument comprising a highly developed closed circuit television system which intensifies light. The detector includes an optical imaging system, image orthicon camera tube and an amplifier. The monitor incorporated several amplifiers with a power supply and a 10 in. kinescope tube on which the intensified image is produced. The imaging system employed an 85 mm, $f/1{\cdot}4$ lens. A timer was mounted adjacent to the kinescope tube face, to record the time of the event on each exposure. A 16 mm movie camera, set at 8 frames/sec, $f/2{\cdot}0$ focused at 41 in. and employing Tri-X film, was used to photograph the kinescope tube.

A bank of six 35 mm cameras equipped with various filters and exposure times was used to record specific features of the cloud luminosity. The cameras were mounted on a single base and pre-pointed so that the cloud would appear in the middle of their field of view which was about 30° in all cases. The items 1 through 6 in Table 1 give the camera type, lens speed, filters, film type, film sensitivity, exposure time and the duration over which usable cloud images were obtained.

3.2. *Optical results*

Easily visible and photographic clouds were generated in each of the two experiments reported here. The optical data obtained are briefly described below, with no analysis presented here except for determination of the initial cloud geometry and location. A more thorough analysis is presented later with the aid of some of the optical data summarized in Table 1.

All three Astro Ballistic camera sites reported positive results which indicated that the cloud position on the morning twilight release of 21 May 1958 was as follows: the altitude of release was 127 km at a position from the launch site of 16·7 km N and 12·7 km W. The apparent diameter of the cloud (exposure 1 min) was 1·4 km. It must be noted that any apparent movement of the cloud through winds, diffusion, shears, etc., would make the apparent image size larger than the actual. Similar data for the 22 May 1958 morning twilight release yielded the position; altitude of 116 km above a region 14 km N and 9·1 km W from the launch site. The photographed apparent dimensions were found to be 0·8 km × 0·9 km. Again, these dimensions are apparent and represent the integrated signal for a period of 1 min.

The Baker–Nunn satellite tracking camera data were obscure. The data on both the 21 and 22 May flights were rendered difficult to interpret due to a severe diffusiveness of the images. Crude estimates of the initial (within the first 10 sec) minimum cloud dimension yield the values 1·0 and 1·2 km on 21 and 22 May, respectively.

The Lumicon records for the 21 May release

VOL 2 1959

Table 1. Summary of Optical Data For Cesium and Sodium Experiments, May 1958

Camera number	Camera type and film	Wavelength region of interest	Filters	Transmission	Film sensitivity	21 May		22 May	
						Exposure (sec)	Duration (sec)	Exposure (sec)	Duration (sec)
1*	Exacta $f/3.5$ Tri-X	Range of Tri-X	None	—	—	1/50	10	1/25	25
2*	Bolsey $f/3.2$ Tri-X	5890 Å	5% at 5890 Å about 70% remainder of spectrum	5% at 5890 Å	—	1/25	10	1/10	60
3*	Exacta $f/2.8$ Kodak 103-0	3300 Å	Interference type $\Delta\lambda\simeq200$Å peak at 3300 Å	20%	~160	2	40	10	Camera failed
4*	Leica $f/2.8$ Tri-X	5890 Å	Interference type $\Delta\lambda\simeq25$ Å peak at 5887 Å	70%	9·8	10	160	10	400
5*	Leica $f/2.8$ Tri-X	4555 Å	$\Delta\lambda\simeq25$ Å peak at 4556 Å	90%	9·5	10	130	10	none
6*	Rival $f/2.8$ High Speed IR	8521 Å	$\Delta\lambda\simeq25$ Å peak at 8517 Å	90%	3·2	10	>200	10	none
7	Beattie-Coleman $f/2.8$ Tri-X	5890 Å	Interference type $\Delta\lambda\simeq25$ Å peak at 5892 Å	70%	9·8	see text	105	see text	240
8	Lumicon $f/1.4$ not applicable	5890 Å	Interference type $\Delta\lambda\simeq25$ Å peak at 5894 Å	70%	—	—	350	—	>700

* Note: Camera numbers 1 through 6 represented the 35 mm multi-camera device.

indicated that the initial geometry of release was that of a sphere with a diameter of about 0·7 km. The Lumicon recorded data for about 350 sec after release time. The data of the 22 May flight was obscure during the initial period, but at about 1 min after release an elliptically shaped cross-section of the cloud was well defined and measured about $2·2 \times 2·9$ km. Photography of the cloud was possible for over 700 sec after release, at which time it became difficult to photograph. Quantitatively, the Lumicon data are somewhat unreliable and it appears that some form of built-in calibration is required for other than qualitative or comparative work.

The 35 mm multi-camera device indicated an initial sphere of average diameter of about 1·2 km which represented the apparent image within a few seconds after release time on 21 May. Similar data for the 22 May flight indicated an initial sphere of diameter approximately 1·1 km.

The Beattie–Coleman Varitron camera obtained photographs of the cloud for a period of 105 sec and 240 sec for the 21 May and 22 May flights, respectively. The apparent initial sizes for the two same respective flights were 1·2 km and 1·1 km. These initial values represent observations made within the first 10 sec following chemical release.

The several data for the optical probes show at least three significant factors: (i) that release was obtained at 128 km and 116 km, respectively, for 21 May and 22 May 1958; (ii) that the feasibility of utilization of twilight resonance radiation for photographic optics is verified; and (iii) that an initial cloud geometry is that of a sphere with an apparent optical diameter of about 1 km.

3.3. *Radar probes*

The radar probes utilized were also employed in previous experiments of this series and have been discussed elsewhere[1], so that only a brief reference to each probe is given here for completeness. The following were employed for the two morning twilight experiments: (a) the Stanford University 70 kW radar which for this experiment began operation on 49·9 Mc frequency subsequently changed when necessary to a lower (23·1 Mc) frequency; (b) the Laredo Smyth forward-scatter net[2, 3] (Laredo Texas–Utah site and Laredo–Nevada site) was employed to check the 201 Mc forward-scatter capability; (c) the *C*-3 ionosonde[2] at WSPG which obtained 1–25 Mc variable frequency returns every 15 sec; (d) the Ft. Huachuca Arizona-site 1 (120 miles north of Ft. Huachuca) and Ft. Huachuca-site 2 (240 miles east) forward-scatter nets[2] were employed. Forward-scatter between these stations was checked on 810 Mc and 13·56 Mc, (e) the Dovap and Doran HAFB–WSPG net[2] transmitting on 36·94 Mc, 73·88 Mc, 132·48 Mc and 264·96 Mc and, (f) the 6-meter and 2-meter radio amateur networks[2] utilizing coded transmissions of 50·0 Mc and 144·01 Mc, respectively. As in the optical case, the radar results obtained are discussed here with analysis given later for comparison to optical data.

3.4. *Radar results*

The radar probes obtained positive results on the 21 May release of cesium, but uniformly obtained negative results for the 22 May flight in which sodium vapor was released at 116 km. Thus, in view of the optical data of 22 May, it is concluded that proper release of sodium vapor was achieved, but no radar-detectable artificial electron cloud was produced. Thus, the following reported positive radar results apply only to the experiment of the morning twilight release of sodium-doped cesium vapor at 128 km on 21 May 1958.

Positive results were obtained for the Laredo Smyth forward-scatter net. On 21 May a signal identified with the ionized cloud was recorded at 116 sec after launch. The Utah site recorded the effect for a period of 2·3 sec while the Nevada station reported a 3 sec duration.

The Ft. Huachuca station 1 and 2 forward-scatter net reported positive results as a short burst on 810 Mc at both stations for about 1 sec.

The radio amateur 6 m net reported a positive signal for over 15 sec commencing at the proper time after launch (about 118 sec). One station receiving a positive result was situated about 100 miles NE of the launch site and received and recorded the 50·0 Mc transmission from Tuscon, Arizona.

The radio amateur 2-m net was successful in producing two positive contacts on the 21 May release. The signals identified with the artificial electron cloud were detected for a period of over 3 sec from both contacts. These contacts were located at 520 miles SE and 90 miles S of the test site.

The 70 kW Stanford radar, situated 15 miles from the launch site, reported saturated echoes associated with the electron cloud on 23·1 Mc for less than 20 sec at about 118 sec after launch and at a slant range of about 191 km.

The Dovap and Doran networks obtained positive results for the 21 May flight. The data for three operating stations are given in Table 2, which includes the station reporting the result, the frequency received, the time duration and recorded signal strength.

The *C*-3 ionospheric sounding facility at White Sands Proving Ground, New Mexico, yielded very useful positive results for the 21 May morning twilight cesium release. The strong echoes associated with the artificial electron cloud persisted for over 20 min after gas release which made possible the determination of the time variation of both the maximum

Table 2. Dovap–Doran Data for Cesium Release, 21 May 1958

Transmitter stations location (miles from launch site)	Frequency (Mc/s)	Power output (W)			
1 mi E	36·94	1200			
1 mi S	73·88	120			
38 mi SW	132·48	1500			
	264·96	200			
Receiver stations location (miles from launch site)	Frequency	Time duration of recorded data		Recorded reflection (average) (μV)	
		20 May	21 May	20 May	21 May
38 mi SW	36·94	60 sec	9 sec	3	2
25 mi SW	73·88	0	2 sec	0	0·1
72 mi NW	132·48	6 sec	2 sec	0·4	0·7
20 mi NW	36·94	30 sec	—	—	—
	73·8	0	—	0	—
	36·94	64 sec	10 sec	2·5	3
	73·8	9 sec	4 sec	0·6	0·8

frequency returns and slant range. Fig. 1 and 2 summarize the ionogram data which are later analyzed in detail.

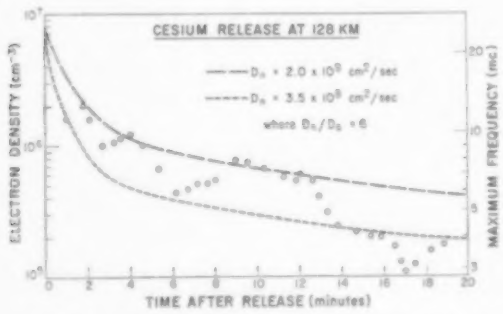


Fig. 1. Plot of the C-3 experimental data for the release of cesium at 128 km. Depicting the best fit values of D_n and D_a. The two closely spaced sets of valves serve to illustrate the sensitivity of the technique.

4.1. *Optical data*

The optical data presented earlier were analyzed and yielded information regarding (a) initial cloud size, (b) height and position by triangulation, (c) diffusion coefficients for the neutral species, (d) the diffusion mechanism, (e) the chemical yield of both cesium and sodium atoms and (f) wind shear values. Items (a) and (b) have been presented earlier and need no further comment, but items (c) through (f) are discussed in detail below.

(i) *Diffusion and yield.* To relate observed or photographed cloud size to the contaminant density and profile throughout the cloud required the use of an expansion model and consideration of the detector response to the cloud luminosity under varying levels of background light. In addition, many factors join

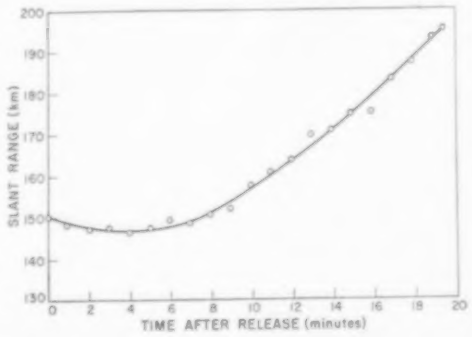


Fig. 2. The circles represent the C–3 radar slant range vs. time for the 21 May experiment. The solid line is a best fit from equation (14) and a constant, horizontal wind velocity of 140 m/sec.

to make quantitative measurements from a photographic image difficult. Some of these are: (1) the maximum exposure time is limited by the background which varies over several decades during the twilight period; (2) the sensitivity of film is affected by the background, becoming more sensitive as the background light brings the level of intensity to a steeper part of the density vs. exposure curve, (3) the optically thin edges of the cloud do not provide a clear cut edge for measurement. In practice, for quantitative work, it is necessary to combine careful calibration and densitometry to plot contour lines of constant image density and relate this to a theoretical model before dimensions can be accurately ascertained. This was not accomplished in the present work. With these limitations stated the model below is employed.

The simplest model is to assume the cloud to be spherical with a Gaussian distribution in density through the cloud. After some initial time the density is expected to vary in time and position in accordance with the equation

$$n(r,t)=\frac{N}{\pi^{3/2}(4Dt+r_0^2)^{3/2}}\exp\left(-\frac{r^2}{4Dt+r_0^2}\right) \quad (1)$$

$n(r,t)$ = number of particles/cm^3 at distance r from center at time t
N = total number of particles in cloud
r_0 = half-width of Gaussian ($1/e$ value) at $t=0$ (cm)
D = diffusion coefficient for neutral or charged species at the cloud height (cm^2/sec).

If luminosity is due to resonance scattered sunlight, and viewed in the back-scattering direction, light will be scattered to the detector in accordance with the number of atoms in a line of sight. The absorption cross-section α, of the sodium atom at its resonance level at 5893 Å is about 10^{-11} cm^2; hence, when the number of atoms in a line of sight is about 10^{11} this portion of the cloud may be considered to be essentially optically dense. At these altitudes collisional deactivation is negligible, so that the optically dense portions have a reflectivity of unity at the characteristic wavelength. Further, the total number of alkali atoms will remain constant since the rate of chemical consumption is very slow. The time for half-consumption is $>10^7$ sec and $>10^5$ sec at 130 and 115 km, respectively (see Fig. 6 of Paper II[21]). τ, the total number of particles in line of sight per cm^2, at a distance from the center r, is given by

$$\tau=\frac{N}{\pi(4Dt+r_0^2)}\exp\left(-\frac{r^2}{4Dt+r_0^2}\right) \quad (2)$$

It is assumed that the optical size of the cloud is determined by the distance from the center d of that isophote for which $\tau=K=1/\alpha$ since the intensity drops rapidly for $\tau<K$; then we have from (2)

$$d^2=-(4Dt+r_0^2)\ln\frac{K\pi}{N}(4Dt+r_0^2) \quad (3)$$

$$d_{\max}{}^2=\frac{N}{e\pi K} \quad (4)$$

$$t_{d_{\max}}=\frac{1}{4D}\left(\frac{N}{e\pi K}-r_0^2\right) \quad (5)$$

$$t_{d=0}=\frac{1}{4D}\left(\frac{N}{\pi K}-r_0^2\right) \quad (6)$$

where $d_{\max}$ = maximum radius to which the isophote corresponding to K atoms/cm^2 grows
$t_{d_{\max}}$ = time to achieve this maximum radius and
$t_{d=0}$ time for cloud to disappear.

It is assumed that the initial, very rapid explosive expansion at release leads to a Gaussian cloud of half-width r_0 and the above mathematical treatment is applicable only subsequent to this initial expansion and it is at this point that we set $t=0$. Since the initial expansion occurs within a second or so, t is essentially the time from release.

It should also be remarked that the above represents minimal treatment of the problem. However, it is sufficient for the present data in view of the limited calibration techniques employed in these initial optical probes. A more comprehensive treatment of this problem will be given in a forthcoming paper.[7]

As an example of the optical model we shall analyze the Beattie–Coleman data for the

21 May flight. Utilizing the optical data, our effort here is to determine the parameters r_0 = the initial Gaussian half-width, N = the number of alkali atoms and D_n = the observed neutral species diffusion coefficient. Similar parameters are later determined in an independent manner from radar data and compared to the optical findings.

From (3) we obtain for $t=0$

$$d_0^2 = -r_0^2 \ln \frac{K\pi r_0^2}{N} \qquad (7)$$

where d_0 = initial optical radius for the isophote corresponding to K. In order to solve (7) for r_0 we need an estimate of N which is obtained from equation (4), since d_{max} is observable. Having obtained r_0 and N we can obtain D from equation (5), since $t_{d_{max}}$ is also an observable. The isophote $K=1/\alpha$ has been selected for this analysis since it is the most easily followed with uncalibrated negatives. The above equations, of course, apply to any isophote.

The values determined above are the first estimate for fitting equation (3) to our data. We subsequently adjust these parameters to obtain the best fit to the entire run of data since the initial values of r_0, D_n and N were obtained from two points, $t=0$, and $t=t_{d_{max}}$.

This fit is surprisingly sensitive since a small variation in either N or D causes large changes in our curve. This is indicated in Fig. 3 where the experimental optical size ($2d$) is plotted against the theoretical curves for two closely spaced values of N.

The fit indicates a value of $D_n = 7\times10^7$ cm^2/sec and $N = 1{\cdot}5\times10^{22}$ atoms or a sodium yield of about 2 per cent. The value of D is consistent with theoretically extrapolated values[8] of D_n at the altitude in question. Furthermore, since $t_{d_{max}} \sim 1/D$ small changes in D shift the maximum of the curve along the t-axis in a sensitive fashion.

A separate determination of the cesium yield has been obtained by observing the disappearance time of its second principal line. When large amounts of a material are released, the cloud expansion can in some cases be studied to greater advantage by observing the growth photographed in the second principal line. This is because such a photographed contour line of constant density per cm^2 column achieves its maximum expansion and its disappearance at earlier time intervals since the value of α is about 100 times less than for the first principal line. The process is thus less effected by uncertainties caused by the changing sky background.

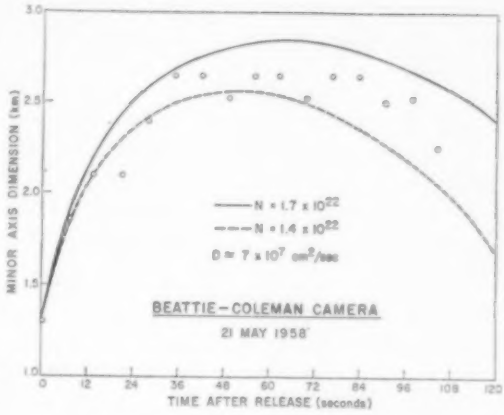


Fig. 3. The time variation of the minor axis optical diameter is indicated by the circles. The dashed and solid curves are theoretical curves for two different yields but the same diffusion coefficient.

Using equation (6) for the disappearance of the contour line defining the optically dense cloud and the observed diffusion coefficient D_n, the yield for cesium vapor for the 21 May 1958 flight was determined. The cloud photographed in 4557 Å lasted for about 130 sec (camera 5, Table 1). If the ratio of optical cross-sections for the first and second resonance lines of cesium is taken to be about 100 to 1, then the recorded contour line corresponded to about 3×10^{13} atoms/cm^2 column. Hence, the total number of cesium atoms released is about 10^{24} or about 3 per cent of the original. This value could not be checked with a similar determination using the photographs in the first line of cesium at 8521 Å since the cloud remained optically dense (in 8521 Å radiation) throughout

the photographic time interval. The photographs in 5890 Å and 3300 Å taken with the six camera bank were also used to determine the sodium yield in a similar manner. These were in agreement with the values established earlier, but were less certain due to the uncertainty in the timing of the individual exposures. However, it is to be noted that in the use of sodium as a "doping" agent for determining yield (by observation of the 5890 Å line), the assumption was made that the chemistry of sodium and cesium are similar. Here, the validity of this assumption is verified since the yields of both cesium and sodium have been ascertained independently and found to agree. These models are applicable only if D_s and the ratio N/K are constant with time.

(ii) *Wind shear.* The previous section has omitted any analysis of the effect of distorting forces on the cloud. But it is evident that distorting forces are present since the photographs show an initially spherical cloud becoming ellipsoidal with time. This effect is attributed to wind shear at the cloud position.

The shear effect may be described by the following equation:

$$\frac{\partial n}{\partial t} = D\nabla^2 n - w\left(\frac{\partial n}{\partial x}\right) \quad \text{where} \tag{8}$$

n = number of particles/cm^3

D = diffusion coefficient, cm^2/sec

w = wind velocity, taken parallel to x-axis, cm/sec.

The treatment of this equation will be included in a forthcoming paper on optical growth[7]. Here, a more crude model is employed but one which gives a good qualitative fit to the data.

The photographed image size of a cloud illuminated by resonance scattering is a complicated function of viewing angle, optical density of the cloud, film sensitivity and exposure time as well as background light intensity. For the purpose of determining shear a simplification is available since it is convenient to consider only the relative values of the minor to the major axis of the elliptical image.

Consider that the cloud in the absence of shear would expand spherically with the diameter $2d_\perp(t)$ identical with the optically observed minor axis. If δ is the coefficient of shear then a differential velocity of the cloud extremes (top edge to bottom edge) is given by

$$v_{diff} = 2\delta d_\perp(t) \tag{9}$$

and the elongation S_h as viewed from below if the upper altitude winds are horizontal is given approximately by

$$S_h = \int v_{diff} dt \tag{10}$$

if the minor axis remains reasonably small compared to the elongation. This simplified model has the advantage of not being sensitively dependent upon the absolute values of the cloud dimensions nor on the isophote (K-value) followed.

The procedure has been to approximate the observed minor axis $2d_\perp(t)$ by a linear expression

$$2d_\perp(t) = a + bt \tag{11}$$

In the case of the 35 mm camera (item 4, Table 1), data for 22 May 1958,

$$2d_\perp(t) = 1{\cdot}4 + \frac{4{\cdot}2}{200}\, t \tag{12}$$

Then the shear produced major axis as observed from below is given by

$$S_h = c_0 + \delta\,(at + \tfrac{1}{2}bt^2) \tag{13}$$

where the constant of integration c_0 is selected to provide a fit for initial values of time.

In Fig. 4 is plotted the variation with time of the minor axis on the left hand graph and on the right hand graph the variation with time of the major axis together with the prediction of equation (13) for data taken with the six camera bank. From these, a value of $\delta = 1{\cdot}1 \times 10^{-2}$ sec^{-1} is obtained. It is assumed in this model that the observed minor axis (lateral expansion) is equal to the vertical expansion.

Similar plots have been applied to the Lumicon data. It is recognized that the absolute cloud dimensions cannot be determined accurately from the Lumicon, due to the numerous changes in the Lumicon sensitivity control during the time observed and due to a

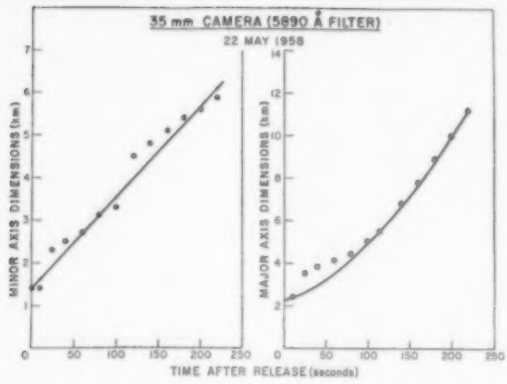


Fig. 4. The solid dots and circles represent the time variation minor and major axes optical diameters, respectively. The two curves show the best theoretical fit and indicate a wind shear of 1×10^{-2} sec^{-1}.

lack of calibration of the photographic film. However, the errors involved are not serious when relative values for the minor and major axis are recorded simultaneously. Data from the 22 May flight yield a shear constant of $1{\cdot}0 \times 10^{-2}\,sec^{-1}$ in agreement with the above while the value obtained for 21 May was $6 \times 10^{-3}\,sec^{-1}$. These values are consistent with those obtained by Manring.[9] It is also recognized that the rate of diffusion will be a function of the cloud shape itself, but this effect is ignored in the above model. Another assumption is that shear does not effect the growth of the minor axis. Yet, despite these limitations a reasonable plot is obtained for the concomitant variation of the major and minor axis.

4.2. *Radar, C-3 data analyses*

Analysis and interpretation of radar data is confined to the *C*-3 results as shown in Figs. 1 and 2, since it is felt that these are the data most amenable for such analysis. Valuable information is obtained concerning diffusion, yield and thermal-ionization efficiency. For this, a more precise mathematical formulation is employed here than that which was utilized in Paper III[3] of this series. In addition, a geometric model is employed to obtain a wind velocity determination at 128 km; this is shown next.

(i) *Wind velocity.* The data shown in Fig. 2 can be utilized for surprisingly accurate wind velocity determinations. The smooth variation of the observed slant range with time, and the character of the curve with a shallow minimum suggest the geometry depicted in Fig. 5,

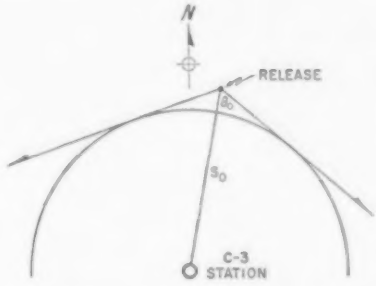


Fig. 5. Schematic representation of geometrical determination of winds from *C*-3 slant rangs vs. time plot.

where the positions of cloud release and *C*-3 stations are included. This model can be described mathematically by the following expression:

$$S = [S_0^2 + vt\,(vt - 2S_0 \cos\theta_0)]^{\frac{1}{2}} \qquad (14)$$

where S_0 = initial slant range

S = slant range

θ_0 = angle between cloud velocity vector and initial position vector from cloud to *C*-3 site

v = wind velocity.

Horizontal winds of constant velocity are assumed.

A best fit to the data was obtained for a v-value of 140 m/sec; the curve is shown as the solid line in Fig. 2. The excellent fit between the experimental points and the theoretical curve seem to corroborate the notion of horizontal winds with constant velocity. The velocity reported here for 128 km is considerably greater than those previously determined in a similar manner, namely, 60 m/sec and 78 m/sec at 95 km[10] and 121 km[3], respectively. Concerning wind direction, it is clear from Fig.

5 that an ambiguity exists on the basis of these data alone. However, visual observations dictated the choice of wind direction as 87 degrees rather than 293 degrees. In summary, the wind determination indicates a horizontal, constant wind velocity at 128 km of 140 m/sec from 87 degrees.

(ii) *Diffusion, yield and ionization efficiency.* In a previous paper[3] a derivation was given for an approximate solution of the mathematical equation which governed the time variation of the radar backscatter frequencies. In order to obtain this solution the ambipolar and neutral diffusion coefficients were equated and the effect of photo-ionization during the rapid expansion phase were neglected. This neglect of photo-electrons during the rapid expansion phase is valid for the case of potassium, but requires more careful analysis for cesium because its photo-ionization probability is over thirty times greater. In the more precise development below we include these effects and limit the solution to the center point solution for the diffusive phase since there is very little variation in density for at least several radar wavelengths in the neighborhood of the center.

It should be repeated that in this paper as in Paper III[3] the only factor causing a decrease in the electron concentration is diffusion. Decay processes such as attachment, recombination, etc., are not sufficiently fast to affect the concentration of electrons significantly for the time constants associated with the effects discussed here.

The governing equations* are given by

$$\frac{\partial n}{\partial t} = D_n \nabla^2 n - kn \tag{15}$$

$$\frac{\partial n_e}{\partial t} = D_a \nabla^2 n_e + kn \tag{16}$$

where n = concentration of neutrals/cm^3
n_e = concentration of electrons/cm^3
D_n = neutral diffusion coefficient, cm^2/sec
D_a = ambipolar diffusion coeffiicent, cm^2/sec
k = photoionization probability per sec.

In order to analyze the initial blast phase we replace the D_n and D_a by some arbitrary D_x representing a forced expansion identical in nature for both ion-electrons and the neutral species. If t_1, represents the duration of the blast we select D_x so that

$$4D_x t_1 = r_0^2 \tag{17}$$

where r_0 represents the Gaussian half-width of the cloud at time t_1. The solution of these modified equations for $t = t_1$ then represent the initial conditions for the second diffusive phase represented by the original unmodified equations, (15) and (16). The blast solutions for $t = t_1$ are

$$n(r, t_1) = \frac{N \exp(-kt_1)}{\pi^{3/2} r_0^3} \exp\left(-\frac{r^2}{r_0^2}\right) \tag{18}$$

$$n_e(r, t_1) = \left\{\frac{N[1 - \exp(-kt_1)] + N_e}{\pi^{3/2} r_0^3}\right\} \exp\left(-\frac{r^2}{r_0^2}\right) \tag{19}$$

where N = the total number of atoms at $t = 0$
$N_e = cN$, total number of electrons at $t = 0$, where
c = thermal ionization efficiency.

All the other symbols remain the same. The expression $N[1 - \exp(-kt_1)]$ represents the photo-electrons. If we put $N = 1 \times 10^{24}$ for the number of cesium atoms as estimated by optical data (and later the fit to the radar curve) and set $N_e = 1 \times 10^{21}$ for a thermal efficiency ionization of $c = 10^{-3}$ as found on previous experiments[3] and on thermo-chemical grounds[4] we find that photo-electrons represent about one-third of the total electrons at $t_1 = 0{\cdot}5$ sec. The estimate of t_1 was experimentally determined from the Lumicon photographs (8 frames/sec) and also agree with the notion of thermal velocities. The above solutions now

* The mathematical solution of the differential equations is due to Dr. Herbert K. Brown, Geophysics Corporation of America.

represent the initial conditions which apply for determining the center point solutions for the normal diffusion phase. These are given for a t'-scale for which $t'=0$ when $t=t_1$.

$$n_e(0,t')=n_{e_0}\frac{r_0^3}{m^{3/2}}+\frac{kN\exp(-kt_1)}{2\pi^{3/2}(D_n-D_a)}\left[\frac{1}{m^{\frac{1}{2}}}-\right.$$
$$\left.-\frac{\exp(-kt')}{p^{1/2}}\right]+\frac{N}{4\pi}\left(\frac{k}{D_n-D_a}\right)^{3/2}\times$$
$$\times\exp\left[\frac{km}{4(D_n-D_a)}\right]\left\{\operatorname{erf}\left[\frac{kp}{4(D_n-D_a)}\right]^{\frac{1}{2}}\right.$$
$$\left.-\operatorname{erf}\left[\frac{km}{4(D_n-D_a)}\right]^{1/2}\right\} \qquad (20)$$

where $m=(r_0^2+4D_at')$, $p=(r_0^2+4D_nt')$, and

$$n_{e_0}=\frac{N[1-\exp(-kt_1)]+N_e}{\pi^{3/2}r_0^3}$$

which represents the initial electron center density at $t'=0$. The plot of this equation is given in Fig. 1 together with the C-3 data; the numerical values of these parameters which gave the best fit were found to be

$$D_n=2{\cdot}5\times10^{8}\ \text{cm}^2/\text{sec}$$
$$D_a=4{\cdot}2\times10^{7}\ \text{cm}^2/\text{sec}$$
$$N=1\times10^{24}\ \text{particles}$$
$$k=6{\cdot}5\times10^{-4}\ \text{sec}^{-1}$$
$$n_{e_0}=7\times10^{6}/\text{cm}^3$$
$$r_0=3\times10^{4}\ \text{cm}$$
$$c=10^{-3}$$

The values of r_0^2 and N were selected on the basis of the optical determinations and were constrained by the necessity of matching the experimental value of n_{e_0}. The value of N_e was determined from thermo-chemical considerations[4] and on the basis of past experimentation.[3] The value of k is considered fixed as discussed previously.[3]

There now remains the problem of evaluating the significance of this model and the parameters determined from it. In the previous model given in Paper III[3], for purposes of mathematical simplicity, no differential diffusion between ion-electrons and neutral species was considered to occur. The solution depended upon selecting a D_e such that $D_a<D_e<D_n$ where D_a and D_n were theoretically determined values and assuming that the charged and neutral species diffused in accordance with an effective D_e. Analogously, a D_e of 9×10^7 cm^2/sec can be fitted to the present data; this value also lies between D_n and D_a. Values obtained above by using the present more sophisticated model, specific allowance is made for the differential diffusion of charged and neutral particles. The present values of D_n and D_a were determined subject to the limitation that $D_n>D_a$ and that the D_n/D_a ratio was to be less than twenty.

The magnitudes of the experimentally determined diffusion coefficients precludes a turbulent diffusion mechanism since the expected turbulence coefficients at this altitude should be orders of magnitude greater. Clearly, no conclusive statement is justified on the basis of this treatment alone. However, an experimental check is available to definitely establish this point. The suggested experiment would involve the twilight release of barium atoms and ions into the upper atmosphere in order to observe their visible resonance radiations so as to determine the individual dynamics of the ion cloud in comparison with that of the neutral cloud.

5. CONCLUSIONS

It is felt that this work has established the value of simultaneous radio-radar and tracer-optical measurements for the determination of the basic parameters and the elucidation of the physical processes active in the growth and decay of artificial electron clouds. There has been developed a simple formula for determining wind which, though it possesses an ambiguity, fits the data adequately. The more sophisticated analysis of the radar data using the independent diffusion of the charged and neutral particles shows some improvement over the simpler model where differential diffusion is ignored. Both the models, one with D_e, the other with D_a and D_n, do not seem adequate to explain the data at $t>800$. Whether this is due to unreliable data or indicates the need of introducing a diffusion coefficient varying with time or shear is not established.

An optical model is employed for the determination of the chemical yield and the diffusion

coefficient. It suffers in this experiment from the lack of calibration but is capable of yielding accurate results.

D_e and D_a (optically) are in fair agreement with the theoretical neutral diffusion coefficient. In addition D_a and D_e determined by radar are also in fair agreement with estimates.

The further development of these experimental techinques and mathematical methods should greatly increase our understanding of artificial electron clouds.

REFERENCES

1. F. F. Marmo, L. Aschenbrand and J. Pressman, *Planet. Space Sci.* **1**, 227 (1959).
2. F. F. Marmo, J. Pressman and L. Aschenbrand, *Planet. Space Sci.* **1**, 291 (1959).
3. F. F. Marmo, L. Aschenbrand and J. Pressman, *Planet. Space Sci.* **1**, 306 (1959).
4. J. Pressman, F. F. Marmo and L. Aschenbrand, *Planet. Space Sci.* **2**, 17 (1959).
5. J. F. Bedinger, E. R. Manring and S. N. Ghosh, *J. Geophys. Res.* **63**, 19 (1958).
6. R. Pendorf, *J. Met.* **4**, 152 (1948).
7. To be published.
8. A. Dalgarno, private communication.
9. E. R. Manring, J. F. Bedinger, H. B. Petit and C. B. Moore, *J. Geophys. Res.* **64**, 587 (1959).
10. F. F. Marmo, J. Pressman, L. Aschenbrand, A. Jursa and M. Zelikoff, *J. Chem. Phys.* **25**, 187 (1956).

VOL
2
1959

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 187–192. Printed in Great Britain.

SIDEREAL ANISOTROPY OF HIGH ENERGY COSMIC RAYS–II

I. ESCOBAR V., N. NERURKAR* and R. WEIL
Laboratorio de Fisica Cosmica de Chacaltaya, U.M.S.A., Bolivia
Centro Brasilero de Pesquisas Fisicas, Brazil

(*Received 5 October* 1959)

Abstract—A report is made of the results obtained from a second period of operation of an extensive shower monitor located at Chacaltaya, Bolivia. The results are consistent with those obtained during the first period of operation, namely, a variation of the order of 1 per cent at 19·0 local Sidereal Time for the high energy showers, and no variation for the lower energy showers recorded. It is shown that the pressure correction applied does not affect these results.

An experiment, monitoring high energy showers has been run at Chacaltaya, Bolivia (5220 m elevation above sea level, 16°S geographic latitude). The detecting equipment consists of four GM trays located at the corners of a 20 m square†. The results obtained over a period of 230 days reported by Escobar *et al.*[1], indicated a diurnal variation with an amplitude of $0{\cdot}88 \pm 0{\cdot}28$ per cent at $16{\cdot}9 \pm 1{\cdot}3$ Local Sidereal Time for showers whose most probable energies were calculated to be 2×10^{15} eV (fourfold coincidences), while for showers of 5×10^{14} eV (threefold coincidences) no appreciable variation was found.

This experiment was continued for an additional period yielding 185 days of useful data.

Figs. 1 and 2 show the solar daily variation of the threefold and fourfold coincidences, respectively, during the first, second and total periods. It can be seen that, considering the statistical deviations, there is a close agreement between the results of the two periods.

The harmonic coefficients for the curves in Figs. 1 and 2 are given in Table 1. The pressure coefficients of the counting rates, calculated for the different periods, are also shown therein. In view of the errors, the coefficients for the threefold and the fourfold rates could be considered as being the same; however, in the total data, which show 0·98 correlation between shower intensity and pressure, the fourfolds appear to have a lower pressure coefficient than the threefolds, in agreement with the previous report.[1]

In Fig. 3 the sidereal daily variation is shown for the threefold rate, both uncorrected and corrected for the contribution of the fourfold rate. The data are consistent with the earlier conclusion that no significant sidereal variation is apparent in the showers responsible for threefold coincidences.

Fig. 4* shows the sidereal variation of the fourfold rate and Table 2 gives the corresponding harmonic coefficients. The dotted curves drawn in this figure represent the first harmonics obtained after all corrections have been applied. No second harmonic contribution is added, as the amplitudes of these harmonics are of the order of the statistical errors.

* On leave of absence from the Physical Research Laboratory, Ahmedabad, India, under UNESCO Technical Assistance Program.

† The complete details of the equipment and its operation are described in *Resumen de Labores* 1956–57, UMSA-CBPF publication. This equipment was loaned by the Laboratory for Nuclear Science of MIT, Cambridge, Mass.

* The results presented in Fig. 4 and Table 2, corresponding to the first period, differ somewhat from those given in Ref. (1). This is due to some computation errors which have been corrected in the present paper.

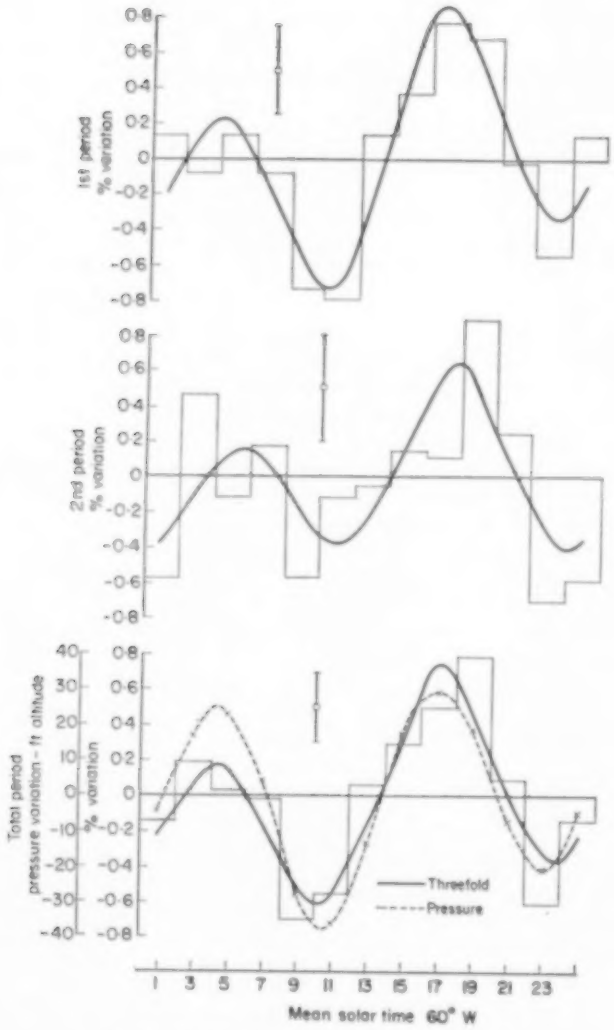


Fig. 1. Diurnal variation of threefold coincidences in solar time without any correction.

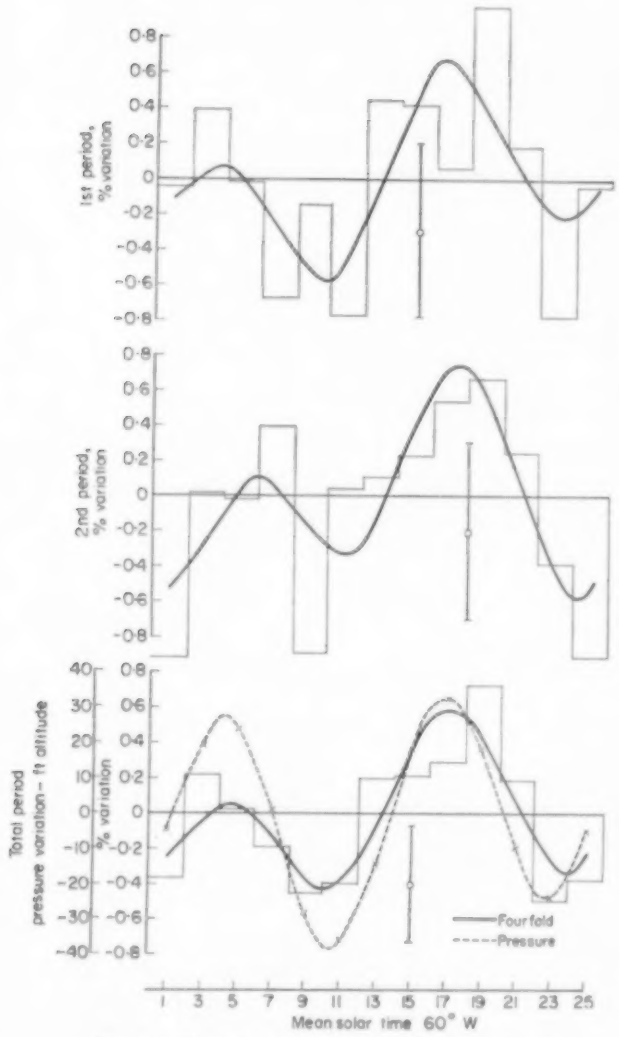


Fig. 2. Diurnal variation of fourfold coincidences in solar time without any correction.

Table 1

Period	Threefolds Solar Time				Pressure threefolds				Pressure coefficient threefolds		
	1st Harm. %	Time	2nd Harm. %	Time	1st Harm. %	Time	2nd Harm. %	Time	*	Correlation	**
1st	0·38±0·11	18·8±1·1	0·51 ±0·11	4·6±0·4	4·5	21·1	28·5	4·9	−11·05±2·38	0·985	−10·75±2·85
2nd	0·29±0·12	17·5±1·6	0·37 ±0·12	5·6±0·6	8·4	21·3	27·7	4·7	− 9·00±3·01	0·900	− 8·22±3·17
total	0·32±0·08	18·7±0·9	0·44 ±0·08	4·9±0·3	6·3	21·0	27·8	4·8	− 9·64±1·68	0·999	−10·35±2·39
Period	Fourfolds Solar Time				Pressure fourfolds				Pressure coefficient fourfolds		
	1st Harm. %	Time	2nd Harm. %	Time	1st Harm. %	Time	2nd Harm. %	Time	*	Correlation	**
1st	0·34±0·19	18·3±2·3	0·34 ±0·19	4·3±1·1	8·0	20·9	29·7	4·7	− 7·08±3·96	0·98	− 7·95±3·70
2nd	0·35±0·20	19·5±2·3	0·40 ±0·20	5·7±1·0	9·5	20·9	28·3	4·7	− 9·96±5·09	0·855	− 8·00±3·75
total	0·29±0·135	18·1±1·9	0·315±0·135	5·0±0·8	8·6	20·9	28·95	4·7	− 6·75±2·89	0·98	− 7·60±2·20

* Pressure coefficient calculated from the 2nd harmonics of intensity and pressure per cent (cm Hg)$^{-1}$

** Pressure coefficient calculated by the least squares method per cent (cm Hg)$^{-1}$

Table 2.

Period	Fourfolds corrected for pressure local Sidereal Time				Fourfolds corrected for pressure "smoothing" and antisidereal (Loc. Sid. Time)			
	1st Harm. %	Time	2nd Harm. %	Time	1st Harm. %	Time	2nd Harm. %	Time
1st	0·40±0·19	17·9±1·5	0·36±0·19	2·1±1·1	0·88±0·28	16·9±1·3	0·45±0·33	1·7±1·5
2nd	0·45±0·20	21·4±1·8	0·27±0·20	2·0±1·6	0·94±0·30	21·6±1·2	0·13±0·37	10·8±6·0
total	0·44±0·13	19·3±1·2	0·31±0·13	2·1±0·9	0·76±0·20	19·0±1·0	0·22±0·23	1·4±6·0

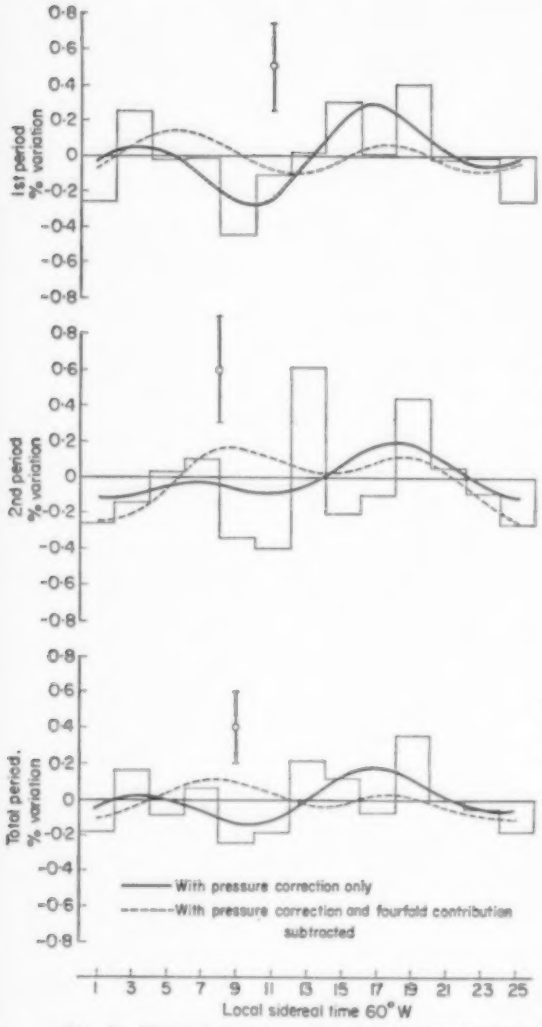

1st period % variation
2nd period % variation
Total period % variation
With pressure correction only
With pressure correction and fourfold contribution subtracted
Local sidereal time 60° W

Fig. 3. Diurnal variation of threefold coincidences in sidereal time.

Fig. 5 shows the sidereal variation of fourfold coincidences with and without pressure correction. There is almost no difference between the two sets of data, so that it is evident that neither the pressure nor the pressure coefficient used have materially affected the sidereal variation found.

It is to be noted that the fourfold sidereal variation in the total data is almost the same as that in the two different periods considered separately, and has not been nullified or reduced to a negligible value as has been the case in results reported by other workers.[2, 3] The time of maximum is 19·0±1·0 Local Sidereal

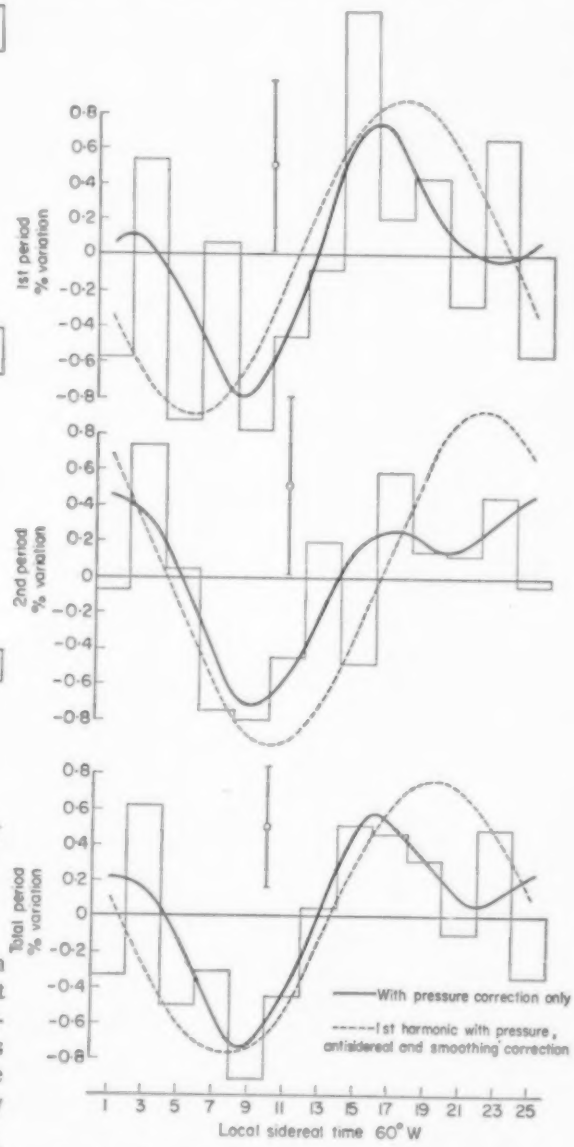

1st period % variation
2nd period % variation
Total period % variation
With pressure correction only
1st harmonic with pressure, antisidereal and smoothing correction
Local sidereal time 60° W

Fig. 4. Diurnal variation of fourfold coincidences in sidereal time.

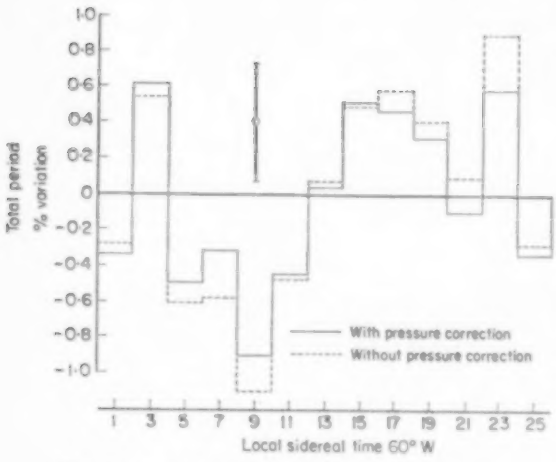


Fig. 5. Fourfold coincidences on sidereal time comparision of data corrected for pressure and uncorrected data for the total period.

Time, which also corresponds closely with the time when the Galactic Center passes closest to the zenith of Chacaltaya.

The consistency of the sidereal variation found is rather unexpected in view of the negative results found by some workers[2, 3] for primaries of about 10^{15} eV. Delvaille *et al.*[4] showed that, in general, sidereal variations reported have the peculiar characteristic of having an amplitude of 3 times the standard error. The results for the two separate periods studied in this paper also come close to this relationship, but it is to be pointed out that the combined data has a somewhat larger ratio of amplitude to error.

CONCLUSION

A sidereal variation is consistently found at Chacaltaya for showers of energies about 2×10^{15} eV. This variation is of the order of $0{\cdot}8 \pm 0{\cdot}2$ per cent and occurs at a time closely approaching the nearest passage of the Galactic Center.

Though this amplitude can be considered significant, being 4 times the standard error, and suggestive of a sidereal anisotropy, it is rather difficult to conclude its existence in view of the arguments of Delvaille *et al.* It is shown that the pressure correction has no effect on the anisotropy found.

Acknowledgements—Our thanks go to Mr. Oscar Saavedra for computational assistance.

REFERENCES

1. I. Escobar, N. Nerurkar, and R. Weil, *Planet. Space Sci.* **1**, 155 (1959).
2. T. E. Cranshaw and W. Galbraith, *Phil. Mag.* **2**, 804 (1957).
3. F. J. M. Farley and J. R. Storey, *Proc. Phys. Soc., Lond.* **B70**, 843 (1957).
4. J. Delvaille, F. Kendziorski and K. Greisen, *Proc. Cosmic Ray Conf. IUPAP, Moscow*, 1959.

BOOKS RECEIVED

KARL KIEPENHEUER: *The Sun.* University of Michigan Press, Ann Arbor, Michigan, 1959. 160 pp.

DONALD H. MENZEL: *Our Sun.* (Revised edition). Harvard University Press, Cambridge, Massachusetts, 1959. 350 pp.

BORIS KIT and DOUGLAS S. EVERED: *Rocket Propellant Handbook.* Macmillan Company, 60 Fifth Avenue, New York, 1960. 354 pp.

PAPERS TO BE PUBLISHED IN FUTURE ISSUES

G. G. BOWMAN: Triple splitting with the *F-2* region of the ionosphere at high and mid-latitudes.

G. G. BOWMAN: Some aspects of sporadic-E at mid-latitudes.

J. PRESSMAN, F. F. MARMO and L. M. ASCHENBRAND: Artificial electron clouds—VI: Low altitude study, release of cesium at 69, 82 and 91 km.

J. CROUCHLEY and N. M. BRICE: A study of "Chorus" observed at Australian stations.

T. OBAYASHI and Y. HAKURA: Polar ionospheric disturbance and solar corpuscular emissions.

A. OMHOLT: The time delay of the red (01) lines in aurorae.

G. J. F. MACDONALD: Stress history of the moon.

S. MIYAMOTO: A geological interpretation of the lunar surface.

A. LUNDBAK: Some simple formulas for latitude effects and lifetime of satellites.

H. NORINDER: Some comments on the penetration of whistlers through ionospheric layers.

L. D. DE FEITER, A. D. FOKKER, H. P. TH. VAN LOHUIZEN and J. ROOSEN: Solar radio events and geomagnetic storms.

L. HARANG and J. TRÖIM: Studies of auroral echoes—I.

M. NICOLET: Report of the COSPAR Space Science Symposium, Nice, January 1960.

V. A. BAILEY: Measurement of the net electric charge on the sun by means of the artificial planet "Pioneer V."

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 195–211. Printed in Great Britain.

SOME ASPECTS OF SPORADIC-*E* AT MID-LATITUDES

G. G. BOWMAN

Physics Department, University of Queensland, Brisbane

(*Received* 16 *September* 1959)

Abstract—Evidence is presented to suggest that the $f_0\mathbf{E}_s$ value tends to be high at positions in the E_s-layer where trough slope-lines and crest slope-lines of *F2*-layer irregularities meet the E_s-layer. These slope-lines are drawn through the troughs and crests, respectively, of the characteristic kinks in the *F2*-layer ionization contours, which are associated with *F2*-layer irregularities.

A rotating-loop direction-finding system has allowed an estimation of the distribution of ionization, which gives rise to Sporadic-*E* echoes. Analysis of Sporadic-*E* occurrence, on two occasions, suggests that the reflecting surfaces are frontal in nature, the fronts having a separation from each other of some tens of kilometres, and probably existing as closed curves, with diameters of the order of several hundreds of kilometres. A possible association between these structures and the occurrence of the green line of the airglow, is discussed.

A distribution of ionization, which will give contours showing "clouds" of ionization at some frequencies and a ripple structure at other frequencies, is proposed, in an endeavour to explain the apparent dual nature of Sporadic-*E* occurrence.

The evidence seems to indicate that the mechanism operating at the E_s-layer level, producing the phenomenon of Sporadic-*E*, is the same as that which produces the *F2*-layer irregularities which are responsible for "Spread-*F*".

1. INTRODUCTION

Many aspects of the Sporadic-*E* phenomenon at a mid-latitude station, Brisbane, have already been described in the literature. (Thomas[2]; McNicol and Gipps[14]).

This paper deals with a relationship between nighttime *F2*-layer irregularities and Sporadic-*E* occurrence, revealed from an analysis of ionograms. A rotating-loop direction-finding system has allowed an estimation of the distribution of ionization at the *E*-layer level, at the time Sporadic-*E* echoes are observed. Rotating-loop records also gave statistics concerning azimuths-of-arrival of signals from E_s-patches, and the lateral deviation of signals, reflected from the E_s-layer, originating at a distant station.

2. EQUIPMENT

Details of the rotating-loop system have been described in another paper[3]. This system was situated at Brisbane, rotated once every 2 min and recorded simultaneously, on separate displays, 3·84 Mc/s transmissions from (i) Armidale (distant 355 km, and bearing 202° from Brisbane), and (ii) a transmitter at the site of the loops. The recording system incorporated a fast swept-gain of 2·5 sec period, and indicated the azimuth-of-arrival of signals received away from the vertical. Records were taken during June, July and August 1958, and for most of the time a 180° ambiguity existed in the recorded azimuth-of-arrival. However, from 3 August 1958, a sensing device was operating which solved the ambiguity on a number of occasions.

3. EXPERIMENTAL RESULTS

(a) *Association with F2-layer phenomena*

F2-layer ionization contours have been constructed from (N, h) curves derived from ionograms by a method due to Schmerling[4]. As an approximation, it has been assumed that all signals are reflected vertically. It has been shown in another paper[3] that the irregularities in the *F2*-layer, responsible for resolved range-spreading, are frontal in nature, their length being of the order of 1000 km. Their cross-

section shows a kinking in the contours of equal ionization, the kink being displaced horizontally with increasing height. Straight lines joining the corresponding points on successive contours have been called "slope-lines". Of particular interest here are the lines joining the troughs and crests, respectively, of the kinks in the contours. The directions of the "slope-lines" seem to be such that they represent a projection of the geomagnetic field lines on the perpendicular cross-section to the *F2*-layer front. The contour shapes produced by this method must be only rough approximations for those irregularities showing a steep rise in the layer level.

When isolated Sporadic-*E* echoes, of a few minutes duration, are recorded they are found to appear close to the points where the trough or crest slope-line of an *F2*-irregularity meets the level of the E_s-layer. Plate 5 of McNicol *et al.*[1] illustrates three examples of E_s-clouds, two of which appear in the vicinity of *F2*-layer resolved range-spreading, while the other has no apparent *F2*-layer association. The variations of ionization contours in the *F2*-layer have been calculated for these three cases, and are shown in Fig. 1. A relationship between the *F2*-layer irregularity and Sporadic-*E* is shown in each case, even though the time displacements between the two phenomena vary considerably. The speed of movement is not known so the plots are not to scale. Fig. 1(a) shows that, at this particular time, the *F2*-layer irregularity existed, but was probably not pronounced enough to produce resolved range-spreading signals. In Plate 5, Fig. 3 of[1], the satellite trace is blanketed at a time when the plot shows the Sporadic-*E* patch to be situated between the *F2* irregularity and the recording station.

A more detailed investigation of the association reveals that, at times, the f_0E_s value varies in such a way that it shows a peak in value at the base of both trough and crest slope-lines of the *F2*-layer irregularity, the crest slope-line f_0E_s peak being less pronounced. Fig. 2(a) and (b) shows this relationship for *F2* irregularities of considerable size, the associated Sporadic-*E* existing for an hour or more. The occurrence of Sporadic-*E* from 00.00 to 02.00, on 27 August 1958, illustrated on Fig. 3(a) is shown, by the *F2*-layer blanketing, to be more intense at 4 points during the period. Fig. 2(c) indicates that these 4 positions appear to be associated with the slope-lines of two small-scale *F2*-layer irregularities.

During the investigation of this *F2*-layer–Sporadic-*E* relationship, Sporadic-*E* patches were found which seemed very likely to be associated with an *F2*-layer irregularity, but occurring earlier than the expected position at the base of the slope-lines. Fig. 4(a) shows an example of this aspect of the phenomenon (Fig. 3(b) is the *h't* record in this case). The

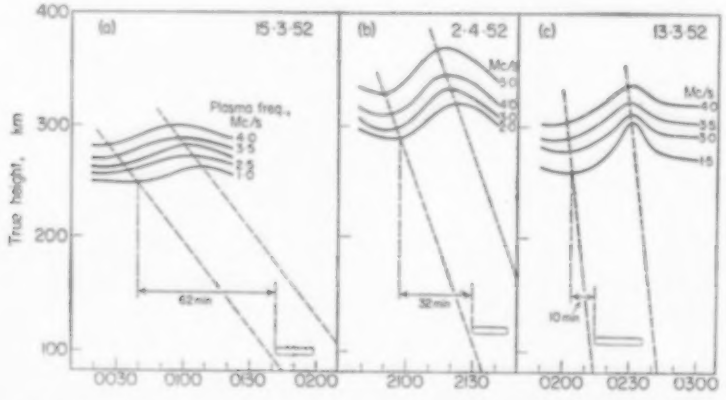


Fig. 1. Association between single patches of Sporadic-E and *F2*-layer irregularities.

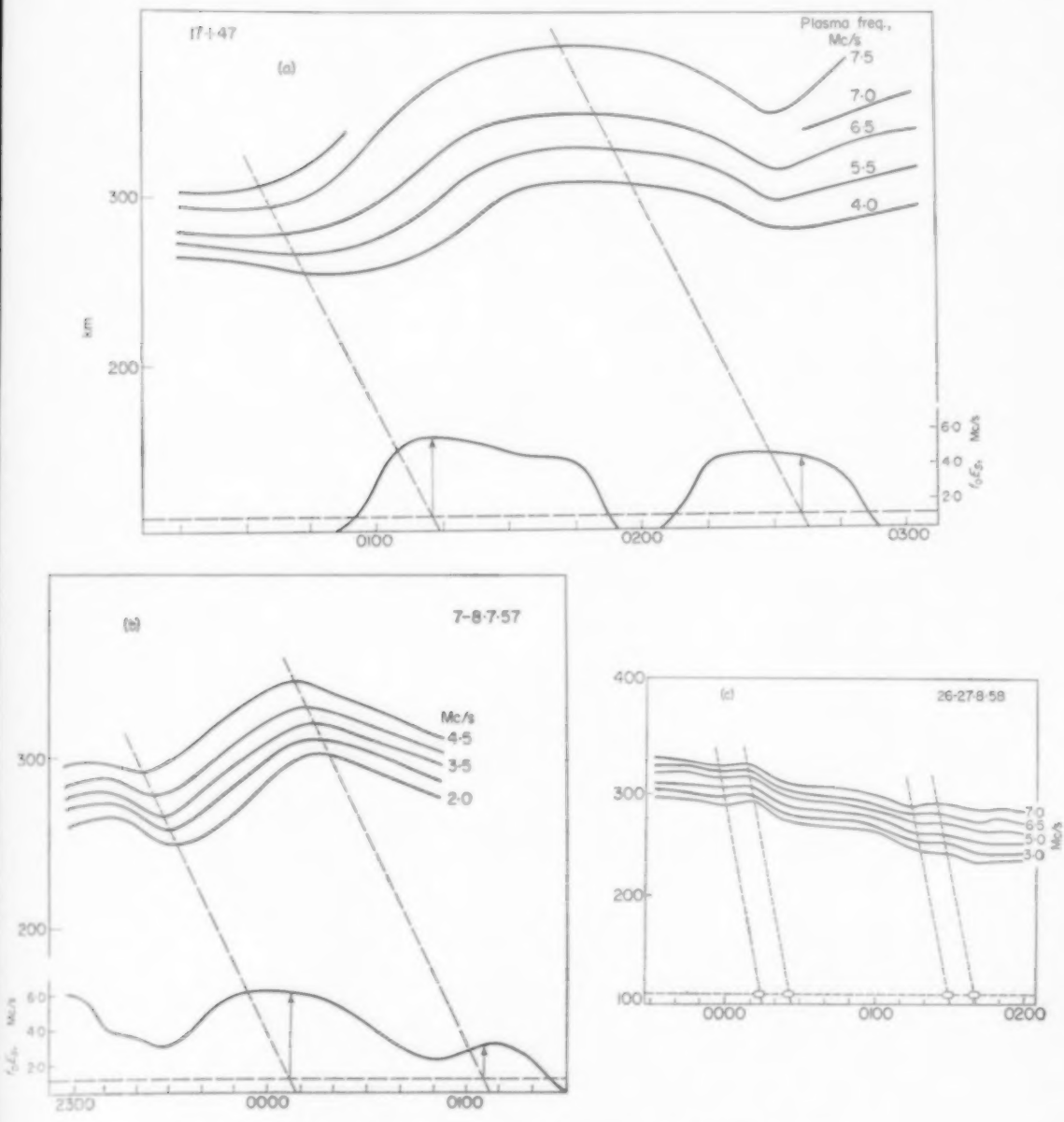


Fig. 2. Association between double patches of Sporadic-*E* and *F2*-layer irregularities.

F2 irregularity is the only large-scale one of the night, and the Sporadic-*E* patch, the first occurrence for 5 hr. In Fig. 4(b), again the *F2* irregularity is the only one of its scale, for the night; also the Sporadic-*E* shown was the only patch observed during the night. The duration

of the *F2*-layer and E_s-layer phenomena are approximately the same here. Fig. 4(c) illustrates that the E_s occurrence is not always early when the *F2* irregularity is a large-scale disturbance with a steep height rise (Fig. 3(c) is the *h't* record in this case).

During the course of work investigating *F2*-layer irregularities, 35 plots were made of contour variation. On twenty occasions Sporadic-*E* was found at or near the slope-lines. These consisted of two cases of a single patch at the foot of the trough slope-line; six cases of a single patch between slope-lines, and twelve cases of an enhancement of f_0E_s at the bottom of both crest and trough slope-lines. Of the other fifteen cases, five had E_s (indicating an

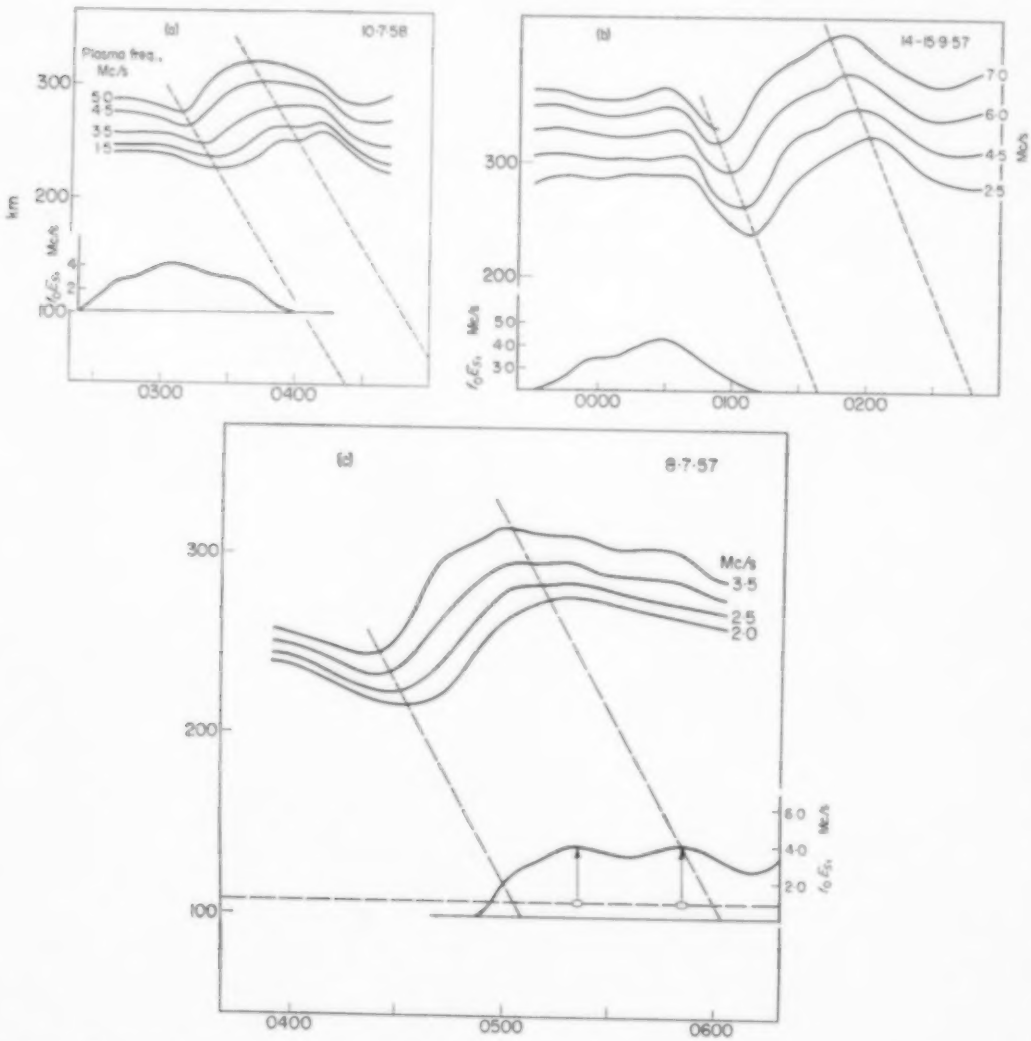


Fig. 4 *F2*-Layer–Sporadic-*E* relationship: (a) and (b) where Sporadic-E occurs before position of slope-lines; (c) *F2*-layer irregularity showing marked height rise with Sporadic-*E* near the slope-lines.

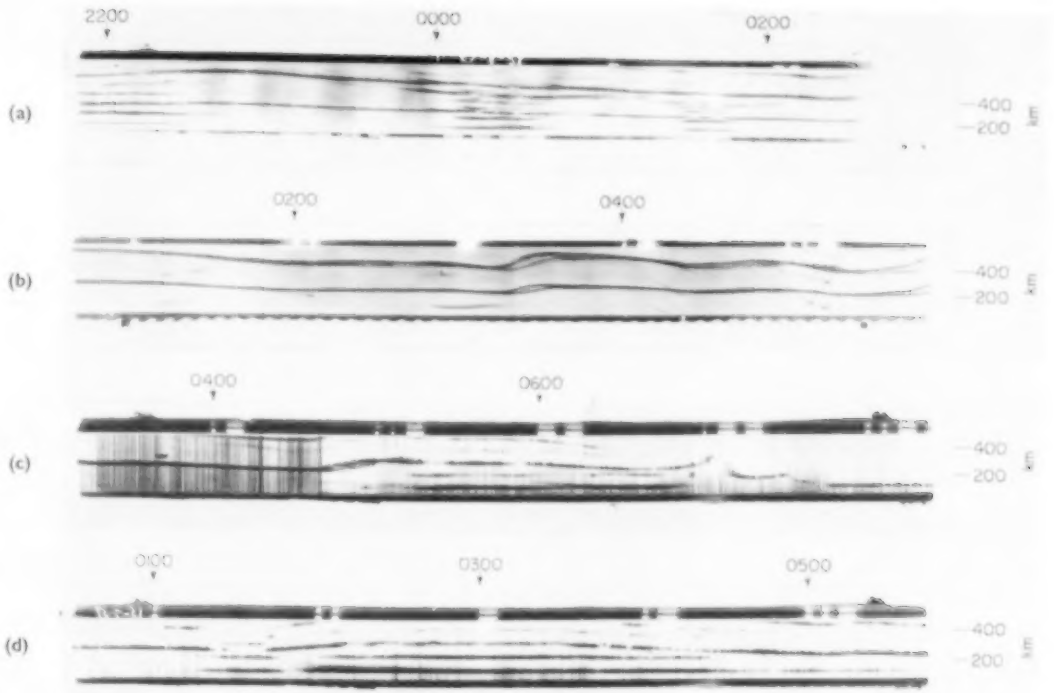


Fig. 3(a). Record ($h't$) at 3·84 Mc s of occurrence of Sporadic-*E* patches in pairs between 00.00 and 02.00 on 27 August 1958. (b) Record ($h't$) at 3·84 Mc s of Sporadic-*E* occurrence prior to *F2*-layer irregularity, detected by lenticel, at about 03.00 on 10 July 1958. (c) Record ($h't$) at 2·28 Mc s of two Sporadic-*E* patches associated with *F2* irregularity at 05.00, 8 July 1957. (d) Record ($h't$) at 2·28 Mc s of Sporadic-*E* occurrence, showing relationship with height changes of *F2*-layer.

facing page 198

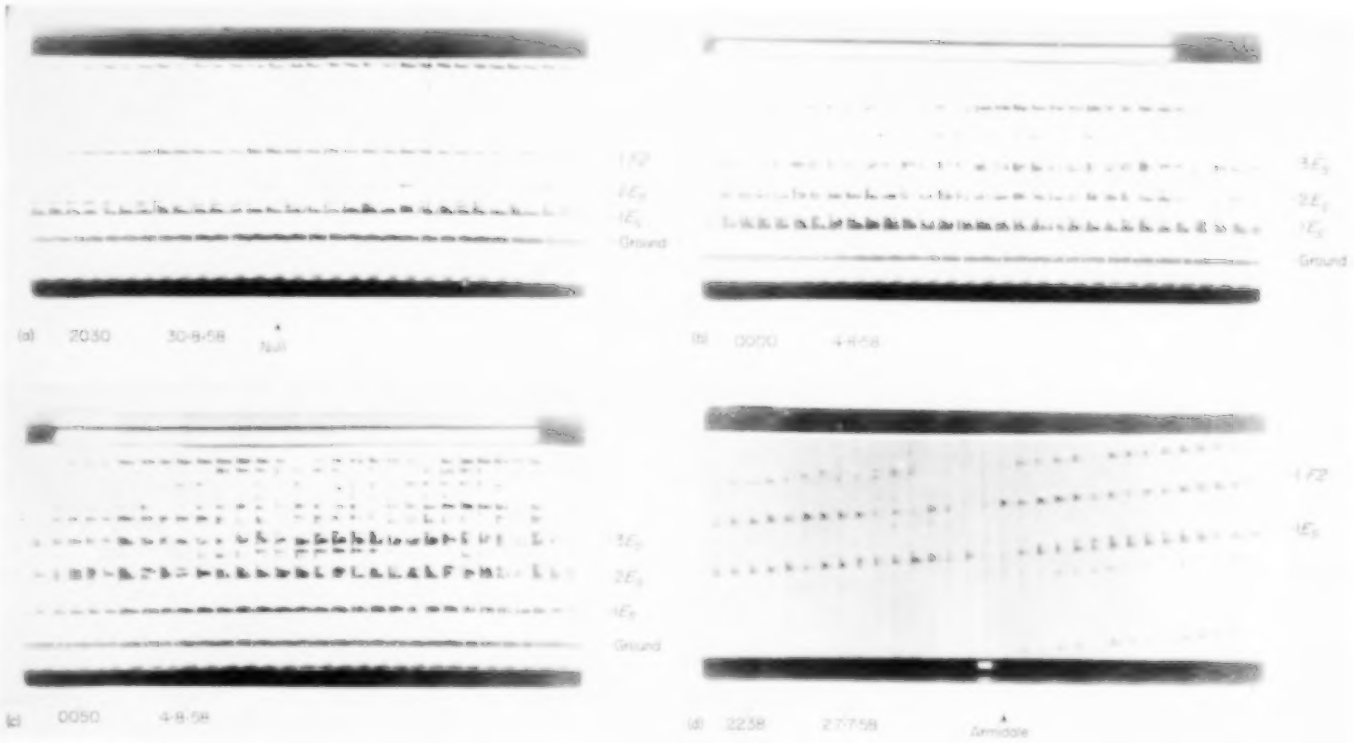


Fig. 5. (a) Rotating-loop normal-incidence record of Sporadic-*E*, showing azimuth-of-arrival null and resolution of satellite traces. (b) Rotating-loop normal-incidence record, showing nature of Sporadic-*E* when spreading appears unresolved. (c) Rotating-loop normal-incidence record, showing satellite signals to two-hop Sporadic-*E* trace, but not to one-hop Sporadic-*E* trace. (d) Rotating-loop oblique-incidence record, showing satellite Sporadic-*E* traces.

apparent association with the *F2*-layer in a manner similar to Fig. 4(a) and (b)) occurring before the position of the slope-lines in the E_s-layer. The remaining ten cases showed no occurrence of E_s in the vicinity of the slope-lines. For the twenty cases where E_s occurs near the slope-lines the average rate of height rise of the *F2*-layer irregularity was 0·67 km/min whereas, for the five cases of E_s occurrence before the slope-lines, the average rate of height rise was 0·88 km/min.

The *F2*-layer–Sporadic-*E* association has been investigated from the other direction. Patches of E_s were chosen because of their isolation from other Sporadic-*E* occurrence, and a plot made of the *F2*-layer ionization contour changes around the times of occurrence of these patches. Thirteen cases were investigated. In eleven cases the patch existed within 10 min of the position of a slope-line from an *F2*-layer irregularity. In the 2 other cases the patches occurred 20 min before the position of slope-lines.

Fig. 2(a) represents the *F2*-layer ionization contour changes of an irregularity responsible for resolved frequency-spreading (cf. Ref. (3) for definitions). Fig. 1(b) illustrates an irregularity associated with resolved range-spreading whereas Fig. 4(c) represents the large-scale ionospheric disturbance reported by Heisler[5] and also reported as lenticels in[1].

The similarity of *F2* irregularity shapes in these cases, and the similar E_s association in each case, suggests that the nature of the ionospheric disturbance, for all three phenomena, is the same.

(b) *Statistics from rotating-loop records*

(i) *Azimuths-of-arrival at normal-incidence.* The normal-incidence recordings of the rotating-loop system gave results indicating azimuths-of-arrival of non-vertical signals for a period from 20 June 1958 to 30 August 1958 (inclusive). Sporadic-*E* traces, showing a directional effect, can be classed into two groups. Azimuths-of-arrival are recorded when a patch of ionization, which has been called an E_s-cloud, approaches or recedes from the recording station. Also, when the Sporadic-*E* trace shows a spreading in range[2] azimuths-of-arrival of the satellite echoes, constituting the spreading, are recorded. When the directional effect lasted for more than an hour, readings have been made every half hour. This method gave forty-three readings during the period of recording, fourteen from E_s-clouds and the remaining twenty-nine from a spreading of the E_s-trace. The fast swept-gain period of 2·5 sec was responsible for resolution of satellite traces in the case of a phenomenon in the *F2*-layer previously called unresolved range-spreading[3]. It has also been effective in the resolution of these diffuse Sporadic-*E* traces in all but five of the twenty-nine cases recorded. Resolution usually revealed three or more satellite traces, and Fig. 5(a) shows an example of this. As in the case of "Spread-*F*", this resolution, and the indication of azimuths-of-arrival for Sporadic-*E* spread echoes, is evidence for a frontal nature for the reflecting surfaces. Five cases were recorded where satellite echoes occurred on the two-hop E_s-trace with little indication of it on the one-hop E_s-trace. This is illustrated in Fig. 5(c). Fig. 5(d) shows one of the few examples of Sporadic-*E* satellite traces recorded at oblique-incidence, and Fig. 5(b) indicates that, when the spreading at normal-incidence is diffuse on rotating-loop records, there is still a suggestion of resolution on some of the swept-gain patches.

Most of the azimuths have an ambiguity of 180°, but a sensing device operated for a short period. The distribution of the 7 cases recorded in this period, where the ambiguity could be resolved, is shown in Fig. 6(e). Fig. 6(b) shows the distribution of E_s-cloud azimuths; Fig. 6(a) the Spread-E_s azimuths and Fig. 6(c), a grouping of Fig. 6(a) and (b); all plottings being made assuming the azimuth-of-arrival to lie between 240° through 360° to 060° (cf.[3]). Fig. 6(d) is a similar plot for the azimuths-of-arrival of "Spread-*F*" signals (a combustion of Figs. 8 and 12 of[3].

(ii) *Azimuths-of-arrival at oblique-incidence.* Azimuths-of-arrival of signals from the distant transmitter at Armidale have been

recorded when they were reflected from the E_s-layer. At times, the azimuth recorded showed a deviation from the true direction of Armidale. Fig. 7(a) shows the distribution of azimuths for readings taken every 6 min, plotted in units of $7\frac{1}{2}°$. The distribution shows a slight preference for westerly deviations. Fig. 7(b) shows the distribution found for similar readings for the first *F2*-reflection when the oblique-incidence record shows no Sporadic-*E* signal. Fig. 7(c) shows the distribution for the same *F2*-trace when Sporadic-*E* signals are present on the oblique-incidence record. In both cases of *F2*-layer signal deviations there is a marked preference for westerly deviations, but particularly when Sporadic-*E* is present, as here the peak of the distribution occurs at $7\frac{1}{2}°$ W.

(c) *Other features of Sporadic-E occurrence*

(i) *Phase-path records.* The nature of these recordings has been described in[1]. An examination of these records shows that, on most occasions, a particular phase-path change, indicated from the record of the E_s-layer, is not superimposed on the phase-path changes of the *F2*-layer. This suggests that phase-path changes in the E_s-layer are not due to changes in the amount of ionization below the E_s-layer, but are possibly due to real changes of range of the reflecting surface. Fig. 8(b) shows an example of an E_s-layer phase-path variation sometimes encountered. The periodic variation suggests a rippled reflecting surface if the changes indicate true range changes. A rippled surface seems likely[3] to be responsible for "Spread-*F*" echoes in the *F2*-layer. True height changes suggested from Fig. 8(b) have been calculated at the turning points of fringe pattern variation, and are shown plotted against time in Fig. 9. If a speed of 240 km/hr is assumed[6] the horizontal dimensions are as shown. The

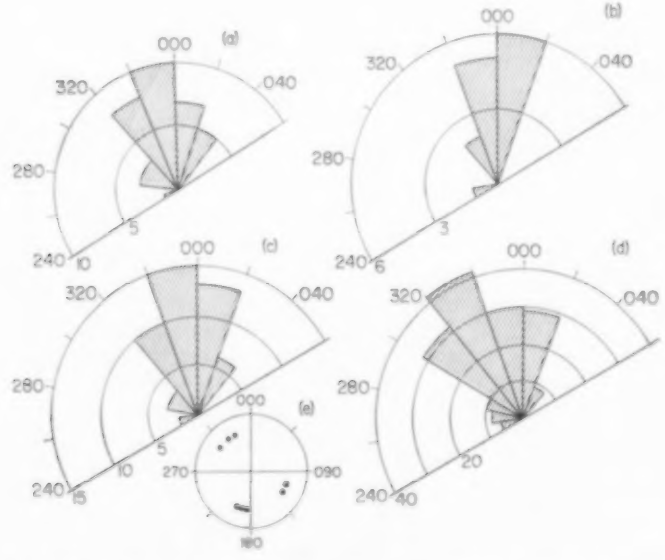


Fig. 6. Distributions for normal-incidence recordings of azimuths-of-arrival. (a) For "Spread-E_s" echoes, assuming azimuths to occur from 240°, through 000° to 060°. (b) For reflections from Sporadic-E clouds, assuming azimuths to occur from 240°, through 000° to 060°. (c) Combination of (a) and (b). (d) For "Spread-*F*" echoes. (e) For Sporadic-E echoes when ambiguity of 180° is resolved.

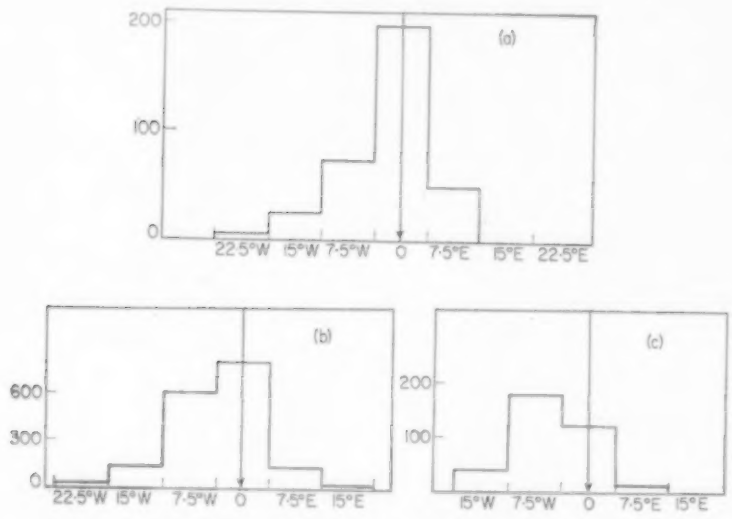


Fig. 7. Oblique-incidence azimuths-of-arrival of signals from Armidale : (a) for signals reflected from Sporadic-*E* ; (b) for signals reflected from *F*2-layer when no oblique-incidence Sporadic-*E* recorded ; (c) for signals reflected from *F*2-layer when oblique-incidence Sporadic-*E* is recorded.

vertical scale has been made 20 times greater than this to illustrate the effect. Plotted in this way, the ripples show a range change from trough to crest of roughly 0·25 km, have a "wavelength" of approximately 40 km, and give reflecting surfaces tilted from the horizontal by about 1°.

Even if this interpretation is wrong, and the phase-path changes are due to ionization drifting across below the E_s-layer, the periodic changes illustrated by Fig. 8(b) are still interesting.

(ii) *Quasi-sinusoidal variation in the azimuth-of-arrival of Armidale signals.* Further evidence of a wrinkled under-surface for the E_s-layer, at certain times, is given by the quasi-sinusoidal variation of the azimuth-of-arrival of

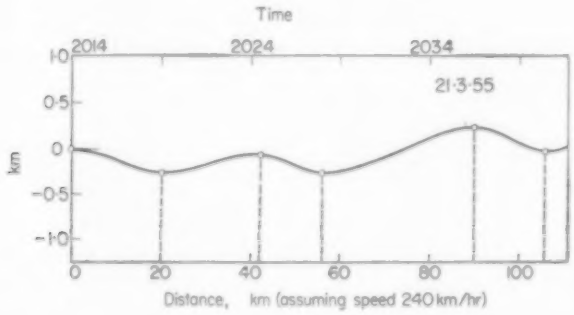


Fig. 9. Variation in height of E_s-layer reflecting surface, if phase-path changes of Fig. 8 (b), are assumed due to real changes in height.

signals, from Armidale, which are reflected from the E_s-layer. This is similar to the effect found for the *F2*-layer[3] but here, in the E_s-layer, the oscillation is usually not as well defined. Fig. 10 shows an example of this variation lasting for about an hour, where each point plotted represents the average of six readings, 1 min apart. During the period of recording, twenty examples of this type of oscillation, for the azimuth-of-arrival of E_s-layer signals, were found.

(iii) *Pairing of E_s-patches.* Sporadic-*E* patches not only tend to occur in pairs related to the one *F2*-layer irregularity (crest slope-line and trough slope-line patches), but also in sets of two apparently associated with different *F2* irregularities. Fig. 3(a) illustrates both types of pairing. July, August and September 1958, 3·84 Mc/s *h't* records have been examined for these phenomena. During this interval fifteen cases were found where the E_s-patches appeared in groups of two separated by time intervals ranging from 30 min to 2 hr. Of these groups, nine cases indicated two groups of two at the operating frequency (3·84 Mc/s), in a manner similar to that shown in Fig. 3(a).

(iv) *Height changes in E_s-patches.* The records suggest that, at times, particularly when the associated *F2* irregularity shows a marked height rise, the height of reflection of the Sporadic-*E* patch related to the crest slope-line is several kilometres higher than that associated with the trough slope-line of the same irregularity. However, this is complicated by the fact that the crest slope-line Sporadic-*E* usually appears to be of lower intensity than the trough slope-line patch[7]. Fig. 3(b) shows the E_s occurrence as consisting of 2 patches, the base of the second one being somewhat higher than that of the first. Fig. 3(c) illustrates the E_s occurrence associated with the *F2* irregularity illustrated in Fig. 4(c). Correction to the observed height for the echo-amplitude of the patch can be calculated from the formula, $\Delta = (8 - 0{\cdot}2W)$ km, where W is the width of the patch in kilometres.[8] The corrections for the two patches of Fig. 3(c), related to the *F2* irregularity, indicate that the height of reflection of the second patch is 1 km lower than recorded, and that the first patch gives the true height of reflection. It is not possible to measure the observed difference accurately to 1 km, so it is difficult to determine whether or not the recorded height differences are instrumental. However Fig. 3(d) shows an example of one patch of E_s being associated with a trough in the 3·84 Mc/s ionization contours of the *F2*-layer, while a more extensive patch seems to be associated with the subsequent *F2*-layer height rise lasting for about two hours. Here the height difference is approximately 10 km, and is real because the relative widths of the respective patches show that the corrections mentioned above will increase the observed difference. The variation of recorded height of the extensive patch of E_s appears to be directly related to the change in height of the *F2*-layer. The true height of the 3·84 Mc/s ionization contour at 01.20 is 268 km, and at 03.50 it is 297 km[4], so that the observed *F2*-layer height change is mainly a real change in the height of the layer at the frequency of recording.

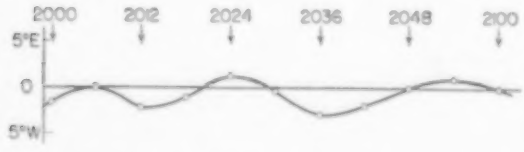


Fig. 10 Quasi-sinusoidal variation of azimuth-of-arrival of signals from Armidale, reflected from E_s-layer, from 20.00 to 21.00 on 10 August, 1958.

(v) *Spread-E_s on ionograms.* While Townsville 1952 ionograms were being examined for "Spread-*F*" phenomenon[3], an interesting feature, relating to the Sporadic-*E* trace, was noticed. When the E_s-trace showed a spreading, this spreading consistently started from a range approximately 40 km in excess of the first E_s-trace recorded. Fig. 8(a) shows an example of this. There is a suggestion of a similar effect at Brisbane, with a 20 km excess range, but it is not well defined, and only occurs occasionally.

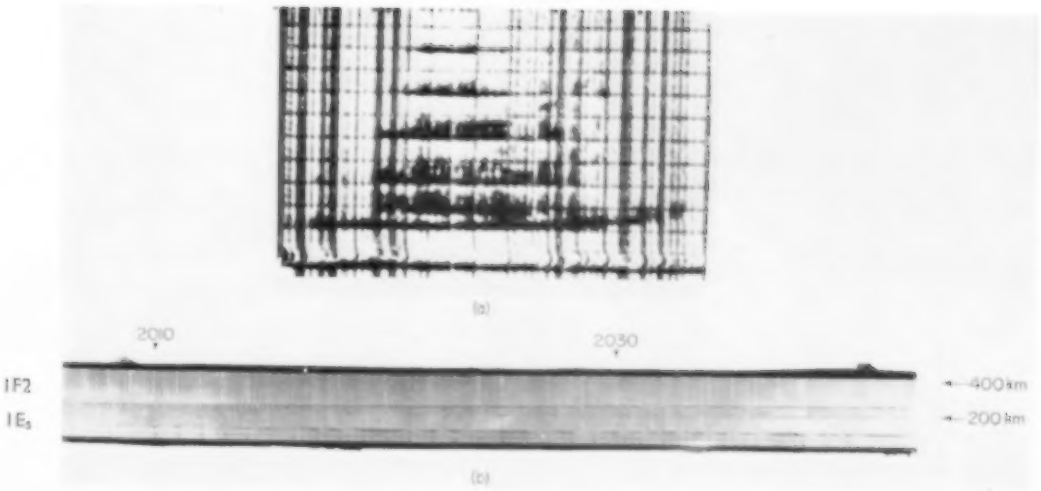


Fig. 8. (a) Ionogram (17.00 on 18 July 1952) showing nature of "Spread-E_s" often seen at Townsville during 1952. (b) Phase-path record (h′t) at 2·28 Mc/s on 21 March 1955, showing systematic phase-path changes in the E_s-layer.

[facing p. 202

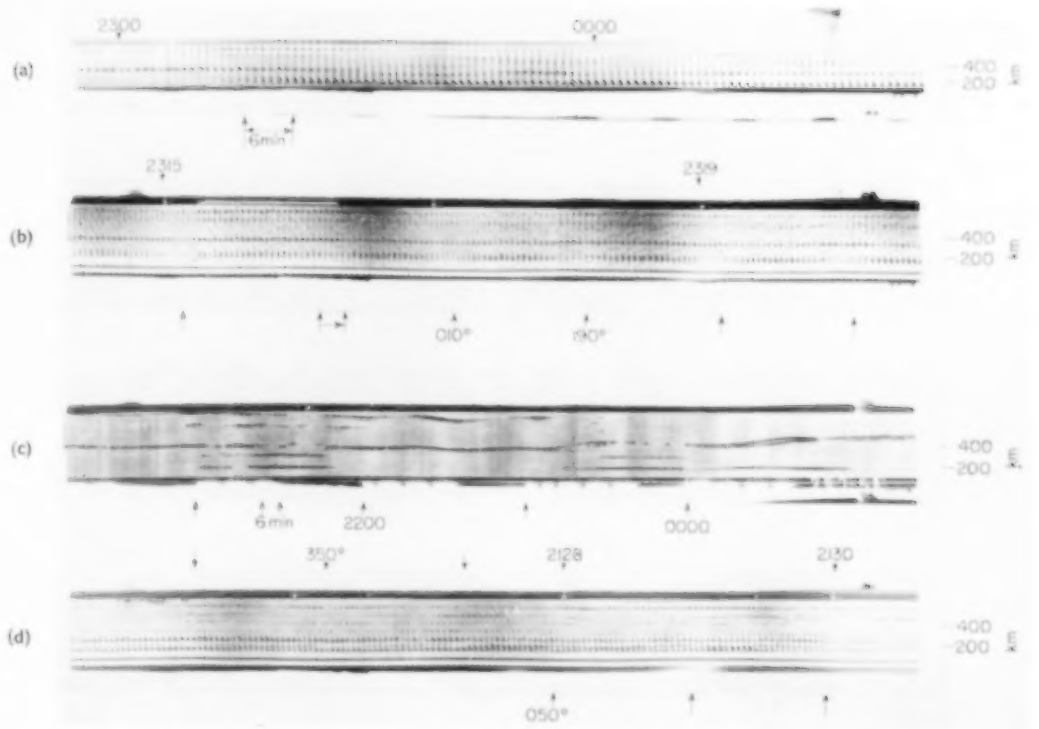


Fig. 11. (a) Record ($h'f$) at 2·28 Mc s of Sporadic-E occurrence between 23.00 on 11 August 1958 and 00.40 on 12 August 1958 with a swept gain period of 1 min. (b) Rotating-loop normal incidence record of Sporadic-E on 11 August 1958, indicating azimuth-of-arrival of Sporadic-E signals. (c) Record ($h'f$) at 3·84 Mc s, showing Sporadic-E occurrence between 20.00 on 27 July 1958 and 01.30 on 28 July 1958. (d) Rotating-loop normal-incidence record on 27 July 1958, indicating a variation in the azimuth-of-arrival of Sporadic-E satellite traces.

(d) *Plan position of Sporadic-E reflecting surfaces*

A special investigation has been made of the occurrence of Sporadic-*E* on the night of 11–12 August 1958. Fig. 11(a) shows the nature of E_s from 23.00 until 00.40 on an *h't* record at 2·28 Mc/s, resulting from a swept-gain method of recording similar to that described in[1]. Here the gain is reduced every minute by 54 dB in nine equal steps of 6 dB. During the periods 23.30–23.45 and 23.57–00.06 the Sporadic-*E* blankets the *F2*-layer, and shows two-hop E_s-traces. Normal-incidence rotating-loop records allowed a determination of the speed of approach of the leading edge of E_s, 23.30 (the time of onset of blanketing) being taken as the time at which the leading edge was overhead. The time when the blanketing ceased has been taken for the overhead position of the receding edge of the disturbance. The speed of approach and the speed of recession have been calculated from the rate of change of range on the rotating-loop record[6]. Fig. 11(b) shows a sample rotating-loop record of the disturbance approaching. The position of nulls for every half-rotation, and the displacement of the null, during the 1-min period when the loops are unbalanced (cf.[3]), indicate that the true azimuth-of-arrival for these signals is 190°. The records also gave the azimuth-of-arrival of signals as the disturbance moved away. The approach velocity was found to be 354 km/hr, 010°, while the recession velocity was 351 km/hr, 350°. It is assumed, at this stage, that reflections are from irregularities, frontal in nature, and that the configuration of the disturbance remains the same for the period being investigated. Further, it is assumed that this Sporadic-*E* structure moves with a constant velocity. The velocity has been determined in Fig. 12 from the two components already calculated, and is found to be 360 km/hr, 000°.

The period of recording on Fig. 11(a) corresponding with the six minutes of recording on Fig. 11(b) is indicated. The fast swept-gain (2·5 sec) of the rotating-loop record reveals that the disturbance gives four reflecting points approaching the recording station. This is not obvious from the 1-min swept-gain records of Fig. 11(a).

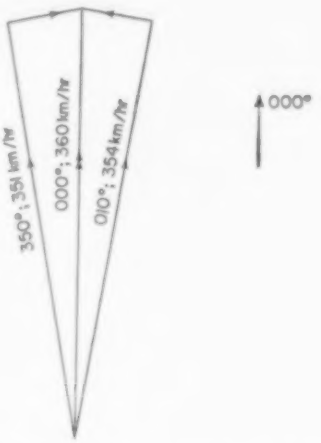


Fig. 12. Velocity of movement of Sporadic-E structure on night of 11–12 August 1958, calculated from two components.

The azimuth-of-arrival of Sporadic-*E* signals was found to vary considerably during the rotating-loop recording on this occasion. Table 1 shows azimuths-of-arrival for one-hop E_s and *M* type E_s echoes,[8] at various times. The horizontal displacements of the reflecting points are also shown when their ranges are well defined in the records. The 180° ambiguity in azimuth readings was not resolved in most cases, and the true azimuths have been assumed as a result of subsequent plotting of the points. No oblique-incidence rotating-loop records were made during the period of this investigation. Fig. 13 shows a plot of the plan position of reflection points; or, when range information is not available, the azimuth-of-arrival is indicated. The configuration is regarded as remaining stationary and the recording station plotting points are moved with the velocity 360 km/hr, 000°. From the information plotted the reflecting surfaces have been estimated, and appear as roughly concentric circles, the outermost circle having a diameter of approximately 200 km.

A detailed investigation of Sporadic-*E* was also made for the night 27–28 July 1958. Fig.

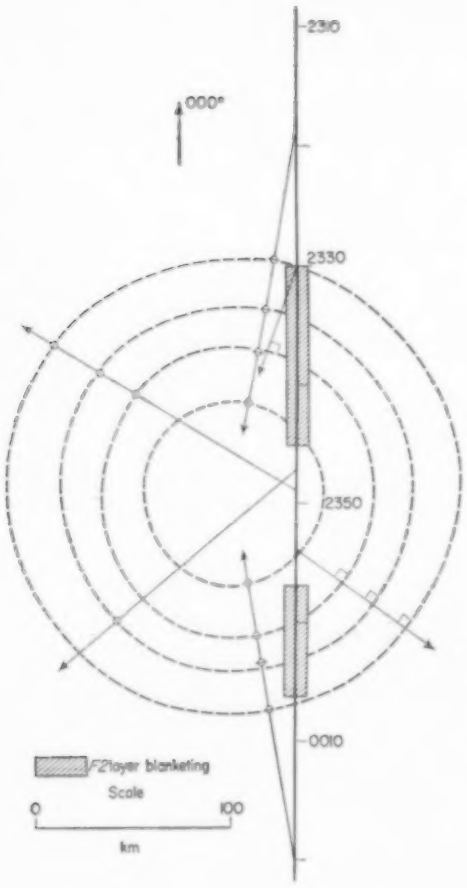


Fig. 13. Sporadic-E reflecting surfaces, estimated from positions of reflecting points for the night of 11–12 August 1958.

14 shows the variation of *F2*-layer ionization contours from 19.00 until 01.00 on this night. Also illustrated is the f_0E_s variation during this period. The f_0E_s value of Sporadic-*E* appears to peak in the vicinity of the slope-lines from two of the large-scale *F2* irregularities, indicated by the ionization contour variation. The same figure illustrates the occurrence of Sporadic-*E* on the oblique-incidence recording on the same night. The times of occurrence at the two reflection positions (vertically above Brisbane and 177·5 km from Brisbane in the direction of Armidale) appear to be related. Fig. 14 also shows times which represent the central positions of blanketing periods indicated on the 2·28 Mc/s and 3·84 Mc/s *h't* records (Fig. 11(c) shows the 3·84 Mc/s information). Between 21.33 and 21.47 on the 3·84 Mc/s record the *F2*-layer is blanketed by the Sporadic-*E*. To determine the position of peaks in ionization density during this period the times (21.36 and 21.45), where the higher multiples of E_s have been recorded, have been used. There is never complete blanketing of the *F2*-trace at oblique-incidence, but points are recorded about which the *F2*-layer signal strength is reduced by more than 20 dB for a few minutes. In Fig. 11(c) the Sporadic-*E* occurrence at about 00.20 (associated with a crest slope-line) appears to be reflected from a higher level than the Sporadic-*E* at 23.30 (associated with a trough slope-line).

Table 1.

Time	Azimuth-of-arrival	Horizontal displacement of reflecting points (km)	Type of reflection
23.18	190°	71 96 118 145	1 E_s
23.30	200°		1 E_s
23.47	230°	120	*M*
23.49	300°	96 118 145	1 E_s
23.54	125°		*M*
00.20	350°	78 102 116 143	1 E_s

Fig. 15 depicts the Sporadic-*E* satellite traces observed from 21.40 to 22.20 on this night. Azimuth readings, from the rotating-loop records, indicate a change in azimuth during the life of this set of divergent satellite traces. This has been interpreted as resulting from a

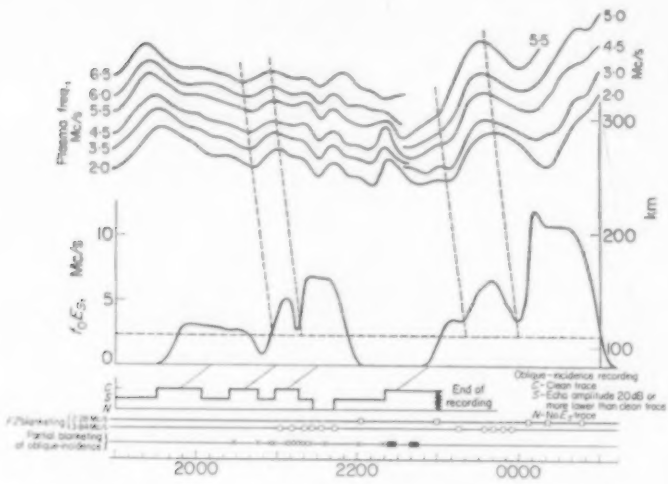


Fig. 14. Ionospheric changes for the night 27–28 July 1958 showing (i) *F2*-layer ionization contour variation; (ii) Sporadic-*E* information at normal and oblique-incidence; (iii) *F2*-layer blanketing information at normal and oblique-incidence.

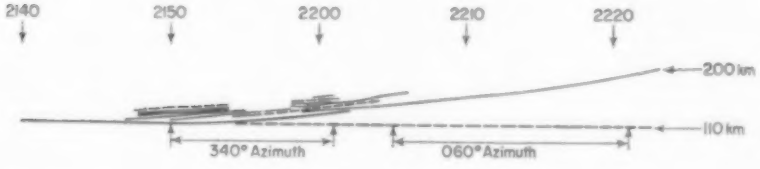


Fig. 15. Illustration of normal-incidence rotating-loop record on 27 July 1958, showing divergent Sporadic-*E* satellite traces between 21.40 and 22.20.

curved front of reflection, similar to that illustrated in Fig. 13. Components of velocity have been calculated, from the rate of change of range, for the two sections of satellite traces where the azimuth-of-arrival remains constant. The true velocity of the reflecting system has been calculated assuming the configuration to retain its shape, and move with a constant velocity. Fig. 16 illustrates this calculation, and gives a velocity of 380 km/hr, 010°.

Fig. 17(a) illustrates a vertical cross-section, through the Brisbane–Armidale line, showing positions (*C* and *E*) where Sporadic-*E* can blanket the *F2*-layer reflection, and the position (*D*) where Sporadic-*E* will give a reflection signal at Brisbane for transmissions from Armidale. In Fig. 17(b) the ionospheric irregularities are regarded as stationary, and the points *B*, *E*, *D*, *C* and *A*, oriented in the Brisbane–Armidale direction, are moved with the velocity of 380 km/hr, 010°. When the positions of strong oblique-incidence Sporadic-*E* signals are plotted along line *D*, it is seen that they occur adjacent to peaks of the f_0E_s value at Brisbane, which occur 25 min later in time. The vertical-incidence blanketing information is plotted along line *B*. Taking into consideration the apparent peaks of Sporadic-*E* ionization plotted along lines *B* and *D*, the oblique-incidence blanketing points have been placed, at the appropriate times, either on line *C* or on line *E*.

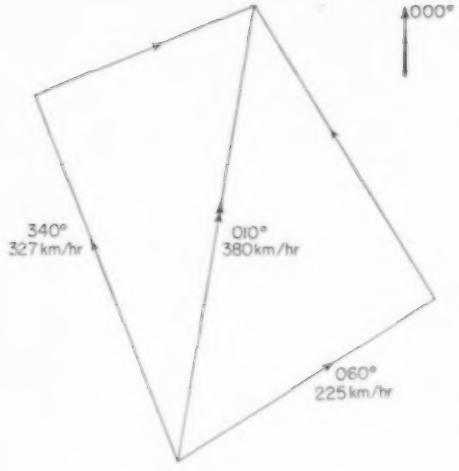


Fig. 16. Velocity of movement of Sporadic-E structure on night of 27–28 July 1958, calculated from two components.

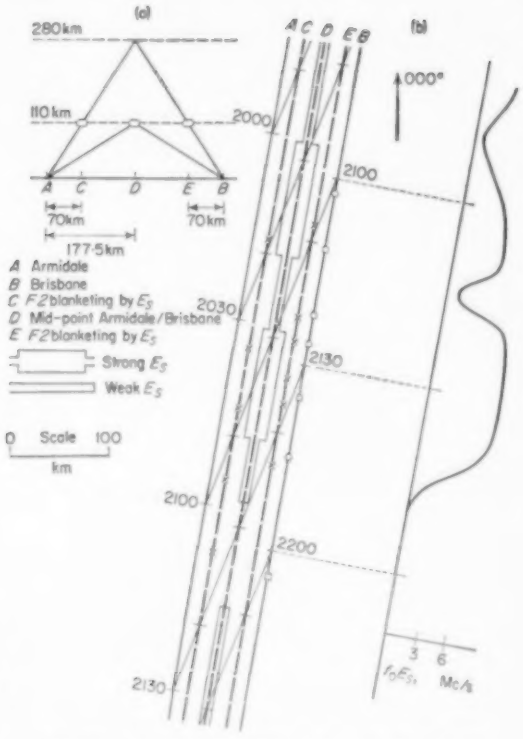


Fig. 17. (a) Vertical cross-section of E_s-and $F2$-layer reflection conditions between Armidale and Brisbane. (b) Plan positions of, (i) Sporadic-E from oblique-incidence reflections, (ii) blanketing Sporadic-E patches recorded at normal and oblique-incidence, from 21.00 to 22.20 on 27 July 1958.

The blanketing information is also shown on Fig. 18. In addition, horizontal displacements of reflection points for satellite E_s echoes at 21.28 and 21.47 are plotted; also plotted are the positions of reflection points used in the calculation of the components of velocity. Fig. 11(d) shows the rotating-loop recording of Sporadic-E around 21.28 and illustrates the changing azimuths-of-arrival for the satellite Sporadic-E signals. The distribution of points Fig. 18 is consistent with the E_s-layer reflecting on structure consisting of a series of fronts, spaced roughly 25 km apart, and having marked curvature. The estimated position of these fronts is shown in Fig. 18; also on a smaller scale in Fig. 19. The rotating-loop recording terminated at 23.02. From 22.54 to 23.02, the normal-incidence recording showed a Sporadic-E trace falling in range from 140 km to 120 km, with an azimuth-of-arrival of 225° (or 045° due to the ambiguity of azimuth). Fig. 19 shows plotted the normal-incidence and oblique-incidence blanketing information from 23.00 to 01.00, and the azimuth-of-arrival of signals from the rotating-loop record of Sporadic-E from 22.54 to 23.02. Points plotted in this figure (Fig. 19) have been used to estimate the positions of reflection fronts passing overhead at Brisbane after 23.00, and these have been related to the fronts found in Fig. 18 by making closed curves of them.

Fronts in the $F2$-layer have been found on this night with a spacing similar to that found here. The positions of these fronts in the $F2$-layer at 20.00 is shown in Fig. 11(b) of[18]. In addition, two-hop $F2$-layer satellite traces at 22.30 indicate an azimuth-of-arrival of 330° or 150°. Four satellite traces are recorded, closely spaced, indicating small-scale ripples having low amplitudes (cf.[3]).

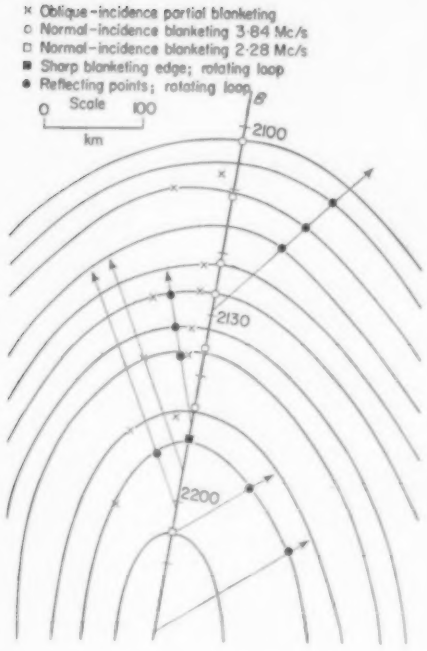


Fig. 18. Sporadic-*E* reflecting surfaces, estimated from the positions of Sporadic-*E* blanketing patches and reflecting points from 21.00 to 22.20 on 27 July 1958.

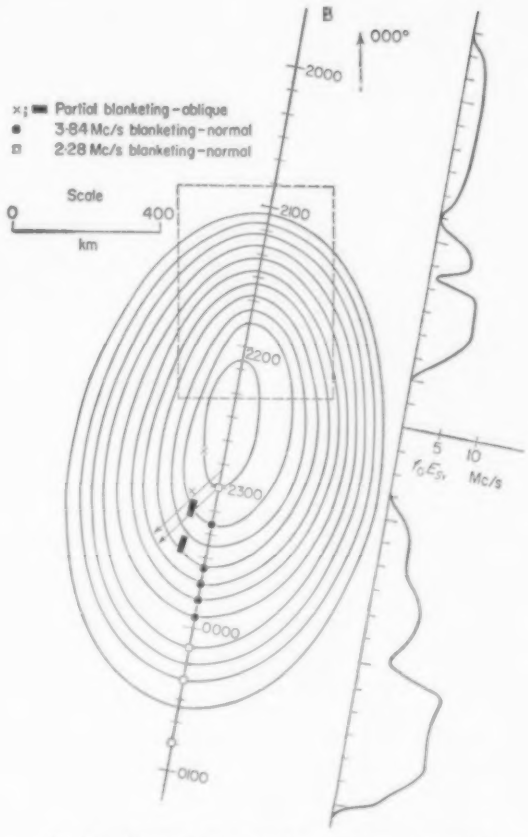


Fig. 19. Sporadic-*E* reflecting surfaces, estimated from the positions of Sporadic-*E* blanketing patches and reflecting points from 21.00 on 27 July 1958 to 01.00 on 28 July 1958.

4. DISCUSSION

(a) *General nature of Sporadic-E from observations*

The results presented in this paper can be summarised as follows. Ionization at the E_s-layer level which is responsible for the phenomenon known as Sporadic-*E* seems to be, on most occasions, denser at positions in the layer down the slope-lines (trough and crest) of the *F2*-layer irregularities. Since the *F2*-layer irregularities appear[3] to be frontal, this implies a similar type of distribution for Sporadic-*E*. Further evidence for this is seen in the distribution of azimuths-of-arrival of Sporadic-*E* traces compared with the corresponding distribution for "Spread-*F*" (Fig. 6). The distributions are similar. Striking independent evidence is revealed in Fig. 7. Here the tendency for the azimuth-of-arrival of the *F2*-layer signal from Armidale to deviate from the true direction is very marked when Sporadic-*E* occurs on the rotating-loop record. This implies that, when Sporadic-*E* exists close to a point immediately beneath the *F2*-layer reflecting point, the probability of a wrinkled *F2*-layer is high. Evidence for a frontal nature in the daytime E_s records has been found by Harvey[9], as on many occasions he records the Sporadic-*E* cloud receding in a different direction to that of its approach.

These results suggest, that, whatever mechanism is operating to cause the characteristic kinks in the *F2*-layer, leading to the phenomenon of "Spread-*F*" described in[3],

this same mechanism is operating simultaneously at the E_s-layer level. The positions of operation in the respective layers seem to be joined, on most occasions, by the geomagnetic lines of force. Each *F2* irregularity tends to give two associated reflecting patches of Sporadic-*E* but as the crest slope-line Sporadic-*E* patch is often weak or missing, a duplication in the E_s-layer, of the number of *F2*-layer satellite traces, may not always occur. The possibility that the crest slope-line Sporadic-*E* patch is a few kilometres higher than the trough slope-line patch (Section 3c(iv)) for large-scale *F2* irregularities also suggests forces acting in the E_s-layer in the same manner as they appear to act in the *F2*-layer. Fig. 20 illustrates the proposed association of reflecting surfaces of satellite traces for the *F2*- and E_s-layers. If this association is true it will not explain the occasional occurrence of Sporadic-*E* a considerable time before the position of the slope-lines in the E_s-layer (Fig. 4), and therefore other factors would seem to be involved in these cases.

The occurrence of E_s-patches in pairs separated by 30 min or more, described in Section 3c(iii), is explained by ionization structures of a nature similar to those of Fig. 13 and Fig. 19.

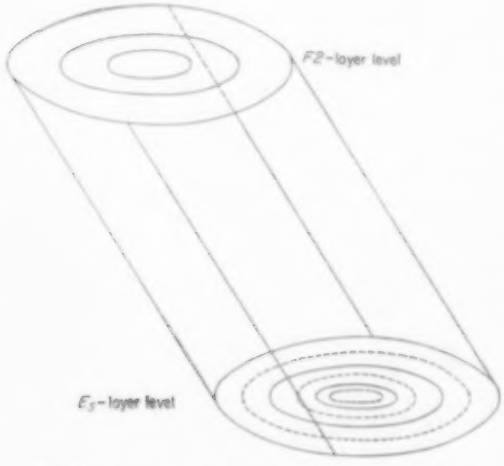


Fig. 20. Proposed relationship between *F2*-layer irregularities and Sporadic-*E* occurrence.

Evidence presented in Section 3c indicates the possibility of the Sporadic-*E* ionization contours being rippled. The phase-path record analysed indicates a rippled surface with ionization contour tilts of the order of 1°. This is consistent with observations by Bramley[10], who found tilts of this order in the Sporadic-*E* layer; also with the tilted Sporadic-*E* surfaces postulated by Thomas[11] to explain double layer phenomena in the E_s-region. Oscillation of the azimuth-of-arrival of oblique-incidence Sporadic-*E* signals, illustrated in Fig. 10, is also indicative of a rippled surface. An explanation for two-hop "Spread-*F*" echoes was made in a previous paper[3] by considering ionization contours, having ripples with amplitudes so small that one-hop satellite traces were not possible. Fig. 5(c) illustrates the same effect for the Sporadic-*E* layer. If a similar explanation is made here, it implies rippled ionization contours. The tendency for westerly deviations for the azimuth-of-arrival of signals from Armidale (Fig. 7(a)) could possibly be caused by one slope of the ripple being steeper than the other, as a result of the role the geomagnetic field might play in the creation of the ripples. Fig. 20 of[3] illustrates this.

A rippled surface for the E_s-layer is in distinct contrast to the ionization patch concept illustrated, for example, by the blanketing information used in Fig. 18. Fig. 21 is an attempt to illustrate a cross-section of the Sporadic-*E* layer ionization contours in the light of evidence presented in this paper. Analogous to what is probably happening to cause *F2*-layer irregularities, forces are visualised which tend to move ionization along the lines of force, acting towards the ground at one part of the disturbance, and in the opposite direction at another. Further if such movements tended to cause a bank-up of ionization near the E_s-layer level, this would give ionization contours at the higher frequencies (e.g. 4 and 5 Mc/s in Fig. 21) which appear as "clouds" and ionization contours at other frequencies (e.g. 3 Mc/s in Fig. 21) which appear as "ripples".

If the Townsville "Spread-E_s" phenomenon,

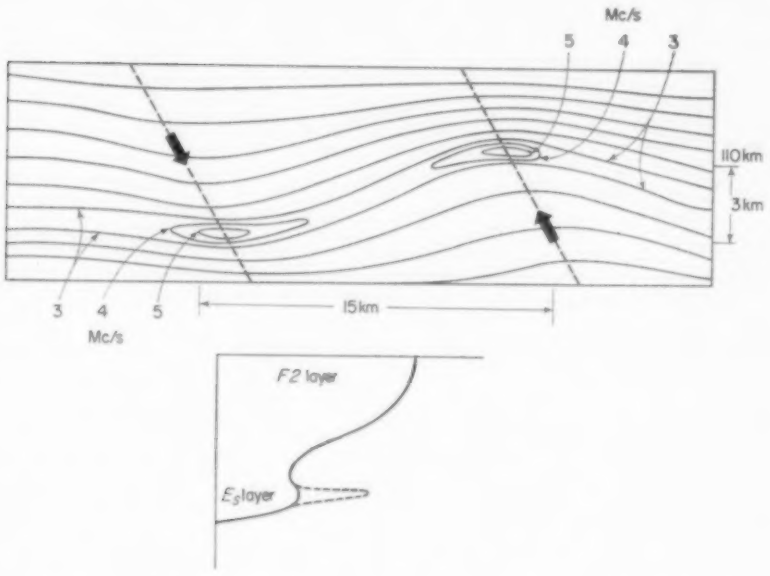


Fig. 21. Proposed distribution of ionization to explain the cloud-like nature and ripple nature of Sporadic-*E* occurrence.

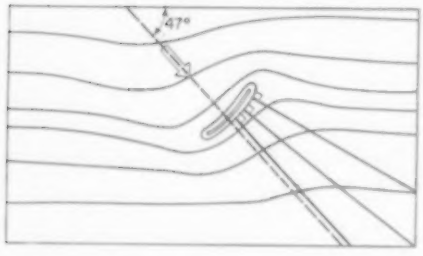


Fig. 22. Proposed distribution of Sporadic-*E* ionization in the plane of the magnetic meridian, to explain some aspects of "Spread-E_s" at Townsville.

illustrated by Fig. 8(a), is associated with fronts having an east-west orientation, the 40-km excess range to the first satellite signal would represent the extra range to the E_s-layer level, if propagation was along the geomagnetic lines of force rather than vertical. Ionization contours, which might possibly explain this aspect of Sporadic-*E*, are shown in Fig. 22, where only satellite signals will be recorded when the angle of elevation of the satellite signal is equal to or less than the dip angle at Townsville (47°).

(b) *Direction of motion of irregularity structure*

The *F2*-layer irregularity structure derived from rotating-loop measurements in[3] appeared to be moving in the direction 010° (magnetic north for Brisbane). This Sporadic-*E* analysis has allowed an estimation of the direction of movement of two Sporadic-*E* irregularity structures, one in the direction of 000° and the other 010°. The results of these few cases suggest that, in general, the direction of movement is close to the direction of magnetic north. Further evidence of this is contained in the paper by Bramley[10]. Here, measurements on a Sporadic-*E* cloud show its velocity as being 50 m/s, 135° for several minutes, when it changed direction and speed to 33 m/s, 225°. If it is assumed, as the results of this present analysis indicate, that the Sporadic-*E* reflections come from frontal irregularities, the true velocity can be deter-

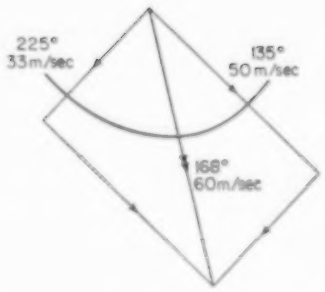


Fig. 23. Movement of Sporadic-E structure, assuming the two velocities recorded by Bramley[10] to be components.

mined from these component velocities. This is found to be 60 m/s, 168° (Fig. 23), which is a movement approximately in the direction of magnetic south, at the locality concerned.

(c) *Association with the green line of the airglow*

Recently[12] airglow cells at the E_s-layer level have been detected, with dimensions comparable with those of the Sporadic-E disturbance derived in Fig. 19. Also, directions of travel towards the equator and speeds of movement[12, 6] are similar for airglow cells and Sporadic-E patches. It is therefore possible that the mechanism responsible for the redistribution of ionization, causing the phenomenon of Sporadic-E, is at the same time responsible for the green line of the airglow.

(d) *Field aligned irregularities*

Aspect-sensitive-reflections along directions perpendicular to the geomagnetic field direction have been attributed to columns of ionization, lying along the geomagnetic lines of force. It is possible that a movement of ionization along the lines of force, as suggested in this paper, will produce not columns but a sheet of concentrated ionization lying along the front of the irregularity at some angle to the vertical. A cross-section of ionization distribution, in the plane of the magnetic meridian, which could possibly give aspect-sensitive-reflections is shown in Fig. 24.

When the cross-section of the front is outside the plane of the magnetic meridian, movement of ionization along the field might be expected to give a slope, which represents the projection of the geomagnetic field direction on this cross-section. Evidence, that the slope-lines of $F2$-layer irregularities obey this relationship, is presented in[3]. Fig. 25 illustrates the relationship. Sloping fronts of this nature in the E_s-layer will give reflections from directions, which are away from the direction of the magnetic pole. The range to these irregularities will be indistinguishable from the range of reflections from columns of ionization (oriented along the field lines), situated away from the magnetic pole direction, when viewed from the recording station.

Acknowledgments—The author would like to thank Professor H. C. Webster for his continued guidance;

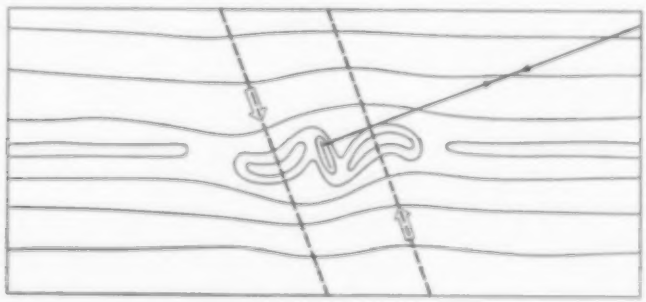


Fig. 24. Proposed distribution of Sporadic-E ionization in the plane of the magnetic meridan giving ionization contours elongated along the geomagnetic lines of force and producing aspect-sensitive reflections.

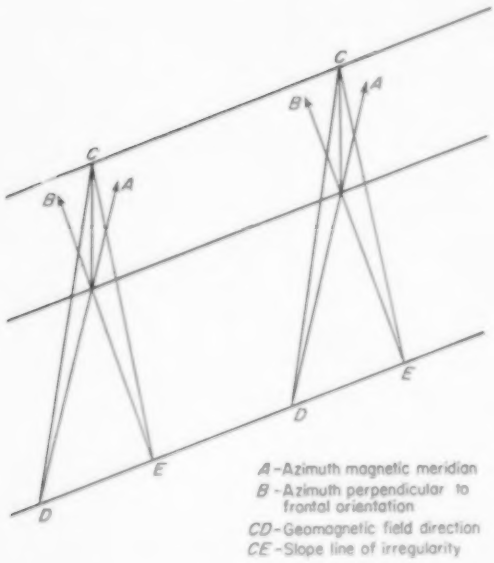


Fig. 25. Relationship between geomagnetic field direction and slope-line orientation of F2-layer irregularities.

also Dr. R. W. E. McNicol and Dr. J. A. Thomas for their assistance during the course of the work. The ionograms were kindly supplied by the Ionospheric Prediction Service and other supplementary ionospheric records used were made available by the Physics Department, University of Queensland. This work was financially supported by the Radio Research Board of C.S.I.R.O. and was carried out while the author was the holder of a C.S.I.R.O. Studentship.

REFERENCES

1. R. W. E. McNicol, H. C. Webster and G. G. Bowman, *Aust. J. Phys.* **9**, 247 (1956).
2. J. A. Thomas, *Aust. J. Phys.* **9**, 228 (1956).
3. G. G. Bowman, *Planet. Space Sci.* **2**, 133 (1960).
4. E. R. Schmerling, *J. Atmos. Terr. Phys.* **12**, 8 (1958).
5. L. H. Heisler, *Nature, Lond.* **183**, 383 (1959).
6. J. A. Thomas and M. J. Burke, *Aust J. Phys.* **9**, 440 (1956).
7. A. J. Lyon and A. J. G. Moorat, *J. Atmos. Terr. Phys.* **8**, 309 (1956).
8. R. H. Hosking, *Aust. J. Phys.* **10**, 220 (1957).
9. J. A. Harvey, *Aust. J. Phys.* **8**, 523 (1955).
10. E. N. Bramley, *Proc. Roy. Soc.* **220**A, 39 (1953).
11. J. A. Thomas, *Aust. J. Phys.* **9**, 574 (1956).
12. F. E. Roach, E. Tandberg-Hanssen and L. R. Megill, *J. Atmos. Terr. Phys.* **13**, 113 (1958).
13. F. E. Roach, E. Tandberg-Hanssen and L. R. Megill, *J. Atmos. Terr. Phys.* **13**, 122 (1958).
14. R. W. E. McNicol and G. DeV. Gipps, *J. Geophys. Res.* **56**, 17 (1951).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 212–213. Printed in Great Britain.

SOME SIMPLE FORMULAS FOR LATITUDE EFFECTS AND LIFETIMES OF SATELLITES

A. LUNDBAK

Danish Meteorological Institute, Geophysical Section

(*Received* 2 *February* 1960)

Abstract—An analysis is made of the effect which manifests itself in the orbital period of a satellite because of the latitudinal shifting of the perigee. The American satellites Discoverer V and Discoverer VI (both in 1959) offered good opportunities for such an analysis, especially after elimination of the normal aperiodic decrease of the orbital period. Hereby, the latitude effect has turned out to mean deflections in the decrease of the orbital period about 20 per cent from the average. Some consequences as to air pressure and lifetimes of satellites are implied.

Normally the perigee of a satellite is the object of gradual and rather regular shiftings in declination (or latitude) and right-ascension. Two kinds of rather slow periodicity in the decline of the orbital period are caused thereby, a day–night effect (the perigee sunlit or not) and a latitude effect.

In 1959 two satellites—1959 epsilon or Discoverer V, launched on 13 August, and 1959 zeta or Discoverer VI, launched on 19 August —offered good opportunities to evaluate the last-mentioned effect. On the other side, the day–night effect was rather limited; moreover, this effect is to some degree eliminated by the following procedure. It may be remembered in this respect that the lifetimes of the two satellites (46 days and 62 days, respectively) occurred around equinox and that the orbital inclinations were 80° and 84°, respectively.

Dealing with some satellite or other it is a well-known complication, that the decline dP/dt in the nodal period P is dependent on the momentary value of P itself. The author has tried to eliminate this aperiodic effect by using the following expression instead of dP/dt:

$$y = \frac{dP}{dt}(P-88)^{2/3} \qquad (1)$$

P is here supposed to be in minutes, whereas the unit of t is taken to be days (also units of longer duration may be chosen of course, for instance when dealing with Vanguard satellites). In our case y has been computed for successive intervals of time (sixteen intervals for Discoverer V and fourteen intervals for Discoverer VI) and plotted against latitude φ of the perigee. Application has been made of the *American Space Track Bulletins*[1] after some control by visual sightings of the satellites, securing a sufficient degree of accuracy for the purpose in question.

During the above plotting the difference between Discoverer V and Discoverer VI appeared to be insignificant, and the following function, which reveals the effect of latitude, could be fitted to the plotted points:

$$y = 0{\cdot}27 + 0{\cdot}06\cos 2\varphi \qquad (2)$$

It is seen that the equatorial value of y is 0·33 and that the polar value of y is 0·21. It must be concluded that the proportion between air drag over the equator and air drag over the pole (in our case the north pole) is just 0·33/0·21 or 1·57, when the distance from the earth's centre is about 6570 km, i.e. the distance of the perigees in question.

Considering that the pole is 21·4 km nearer to the earth's centre than a point on the equator,[2] the above factor 1·57 is fully explained by a scale height of 47·5 km, meaning for instance that the atmospheric pressure diminishes by 50 per cent when the altitude increases by 33 km. Now, at an altitude of

about 200 km the scale height has just this order of magnitude.[3] Possibly the exact value of the scale height at this altitude is a little less[4] with the consequence that the atmospheric pressure at the altitude in question is slightly less over equatorial regions than over polar regions. It may be mentioned in this connection that the atmospheric pressure and the density at an altitude of 175–200 km are 10^{-9} times their values at the earth's surface.[5]

Equation (1) has an important consequence. By integration it can be transformed into

$$T = \frac{3}{5y}(P-88)^{5/3} \qquad (3)$$

where T is the time elapsing from the moment, when the nodal period is P min, until the moment when the nodal period is 88 min, i.e. until the "death" of the satellite. It is seen that the lifetime of a satellite can be computed rather simply, when both P and y (actually some mean value of y) are known. Taking 1958 delta two (Sputnik III) for instance, the relation may be used in the analogous way as equations (1) and (2), y and the two constants in equation (2) being nearly thirds of the corresponding Discoverer-values. An analysis of the Sputnik III-values is in operation but has not yet been brought to an end. It is hoped, moreover, that more extensive analyses will result in a general day–night factor in the y-quantity.

Definitions published on earlier occasions, for instance by H. B. Wackernagel[6], also yield information about lifetimes of satellites; but their starting points and their applicability differ from the formula presented here.

Acknowledgement—It is gratefully acknowledged that the bulk of the computations in question has been performed by Mogens Jörgensen Lyngby State College.

REFERENCES

1. *Space Track Bulletins*, issued by Air Force Cambridge Research Center, Bedford, Massachusetts.
2. D. G. King-Hele and R. H. Merson, *Nature, Lond.* **183**, 881 (1959).
3. B. R. May and D. E. Smith, *Nature, Lond.* **184**, 765 (1959).
4. H. K. Paetzold, *Planet. Space Sci.* **1**, 115 (1959).
5. D. G. King-Hele, *Nature, Lond.* **183**, 1224 (1959).
6. H. B. Wackernagel, *Project Space Track*, AFCRC-TN-59-455 (May 1959).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 214–222. Printed in Great Britain.

TRIPLE SPLITTING WITH THE *F2*-REGION OF THE IONOSPHERE AT HIGH AND MID-LATITUDES

G. G. BOWMAN

Physics Department, University of Queensland, Brisbane

(*Received* 11 *November* 1959)

Abstract—The occurrence of the third (*z*-ray) component of the *F2*-trace on ionograms is investigated at high- and mid-latitudes. Diurnal variations show a systematic shift, with magnetic inclination, of the time of maximum occurrence. Seasonal variations show a winter maximum, and an inverse sunspot-cycle relationship exists. Maximum occurrence appears between a magnetic inclination of 70° and 80° with a fall-off either side.

Evidence is presented to suggest a *z*-ray association with "Spread-*F*" fronts, and a possible mechanism for the recording of the *z*-ray trace at the transmitter site is described. This involves longitudinal propagation of the *o*-mode at its normal reflection level, coupling at this point, and ultimate reflection for the *z*-ray mode as a result of sloping ionization contours belonging to "Spread-*F*" fronts extending in directions perpendicular to the magnetic meridian.

An association with V.L.F. emissions ("dawn-chorus") is discussed.

1. INTRODUCTION

The various hypotheses, which have been advanced explaining the third (*z*-ray) component of the *F2*-trace on ionograms, have been summarised and discussed by other workers[1, 2]. These include (i) coupling, with vertical propagation[3, 4] and (ii) propagation along magnetic field lines, the return of the *z*-ray in this latter case being due to scattering[5]. The former hypothesis explains satisfactorily the observations at geomagnetic latitudes greater than 70°, but only the latter seems capable of explaining the *z*-rays occasionally seen at mid- and low-latitudes[2, 6].

An important contribution was made by Ellis[7], who measured the direction-of-arrival of *z*-rays at Hobart (geomagnetic latitude 51°S) around the sunset period. He found that the observed direction-of-arrival of the *z*-ray component was in the magnetic meridian, the zenith angle being consistent with a ray-path tangential to the magnetic field direction at the normal *o*-ray reflection level, assuming the contours of equal ionization density remained horizontal up to this level. Under such circumstances, coupling should lead to partial transmission (in the *x*-mode) to higher levels, reflection finally occurring at the upper *X*-ray reflection level. Ellis postulates that, in addition to specular reflection (giving a ray which fails to penetrate the lower levels), scattering by small-scale irregularities, in this vicinity, gives a ray retracing its path. Banerji[8] proposed a similar theory, without specifying a mechanism for the return of the energy.

The present paper summarises results from published data for some high-latitude stations, and also gives the statistics (derived from the original ionograms) of *z*-ray occurrence at Hobart for 5 months of 1946. Fifteen years of ionograms (1944 to 1958 inclusive) have been examined for a mid-latitude station, Brisbane, for this phenomenon. During this period 313 cases were found where a third trace exists on the ionograms, spaced in critical frequency from the *o*-ray approximately one-half the electron gyro-frequency. In the phenomenon of resolved frequency-spreading (*vide* Plate 2, Fig. 2, of [9]) it is possible that, occasionally, the satellite *o*-ray trace may be spaced by this interval from the main *o*-ray trace. However, if such a satellite exists, an *X*-ray satellite should also be visible, at least on some part of the record. This was not found for any of the

Table 1.

Station	Geographic coordinates	Geomagnetic latitude	Magnetic inclination
Clyde River	70·5°N, 68·6°W	72°N	84°
Churchill	58·8°N, 94·2°W	68°N	84°
Tiksi Bay	71·6°N, 128·9°E	61°N	80°
Macquarie Is.	54·5°S, 159·0°E	61°S	78°
Hobart	42·9°S, 147·2°E	51°S	72°
Brisbane	27·5°S, 152·9°E	35°S	57°

cases cited. Other evidence also supports their identification as *z*-rays.

The statistics of these occurrences are analysed in Section 2. In addition certain cases have been examined in detail, *F2*-layer ionization contours having been derived from ionograms obtained near the time of the *z*-trace occurrences.

2. RESULTS OF ANALYSIS

(a) *z-Rays at high-latitude stations*

(i) *Introduction.* Geographic co-ordinates, geomagnetic latitudes and magnetic inclinations at the five stations, used in this section, are contained in Table 1. Similar figures for Brisbane (a mid-latitude station) are also included in this table. Results for Churchill and Clyde have been taken from Scott[5] and those for Tiksi Bay from Besprozvannaya and Lovcova[10]. Macquarie Island results come from the A.N.A.R.E. reports[16], while the Hobart figures have been taken from the ionograms.

(ii) *Diurnal variations.* Fig. 1 shows the diurnal variation of *z*-ray occurrence in 1946 at Churchill, Tiksi Bay and Hobart. In the same figure the Macquarie Island diurnal variation is shown for 1951 (a year when the average sunspot number is not far removed from the average sunspot number for 1946). The local time of occurrence of the principal peak seems to undergo a systematic change with the magnetic inclination of the station, from about midday at Churchill (84°) to 06.00 hr at Hobart (72°).

(iii) *Seasonal variations.* For Macquarie Island and Hobart, the seasonal variation of

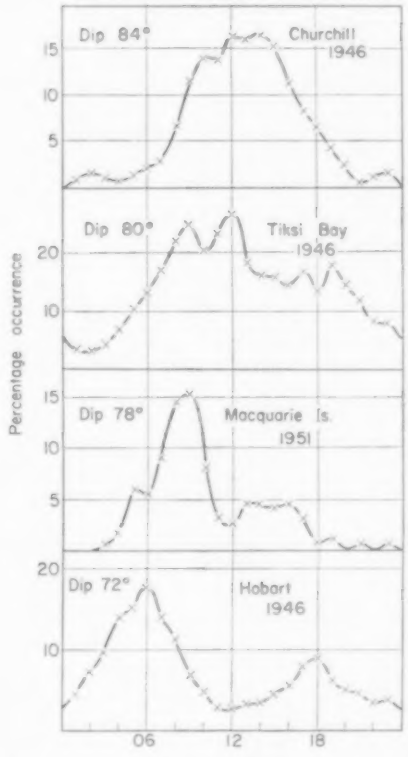


Fig. 1. Diurnal variations of percentage occurrence of *z*-rays at Churchill, Tiksi Bay and Hobart in 1946, and Macquarie Island in 1951.

z-rays is shown in Fig. 2, revealing a winter maximum and a summer minimum in both cases.

(iv) *Sunspot-cycle variation.* Information available about the sunspot-cycle variation of *z*-rays at Churchill, Tiksi Bay and Macquarie

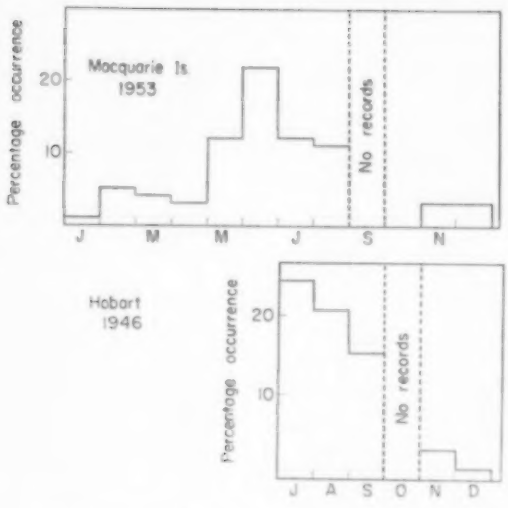


Fig. 2. Seasonal variations of percentage occurrence of z-rays at Macquarie Island (1953) and Hobart (1946).

Island is shown plotted in Fig. 3. All three stations show an inverse relationship with sunspot activity.

(v) *Variation with magnetic inclination.* The percentage occurrence of z-rays at Clyde, Churchill, Tiksi Bay and Hobart during 1946 is shown plotted against magnetic inclination in Fig. 4. The occurrence during 1946 at Brisbane (fifty-eight cases on ionograms taken at 15 min intervals) was so small that its displacement from zero cannot be perceived in the diagram.

(vi) *Nature of triple-splitting at Hobart.* In the majority of cases the z-ray at Hobart appears on an ionogram as a clean single-line trace. But on some occasions this trace is duplicated in the same way as the *o*- and *X*-ray traces are duplicated (resolved frequency-spreading). At other times it apparently consists of a number of traces so close that the z-ray trace appears diffuse. Of the Hobart ionograms for September 1946, on which z-ray traces appear, the z-ray appeared clean on 82 per cent; 2 per cent show the z-ray trace duplicated while on 16 per cent the z-ray appears diffuse. For 9 per cent of these ionograms the *o*- and *X*-ray traces appear as clean traces, resolved spreading is present on 4 per cent, while 87 per cent show unresolved spreading.

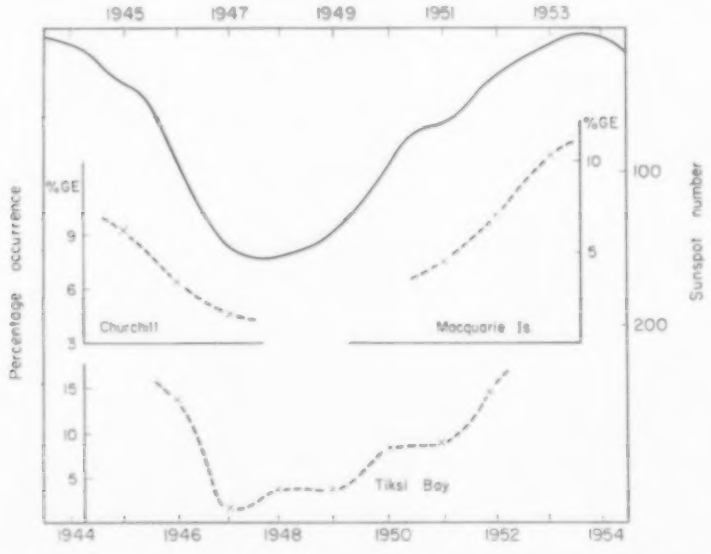


Fig. 3. Sunspot-cycle variations of percentage occurrence of z-rays at Churchill, Tiksi Bay and Macquarie Island.

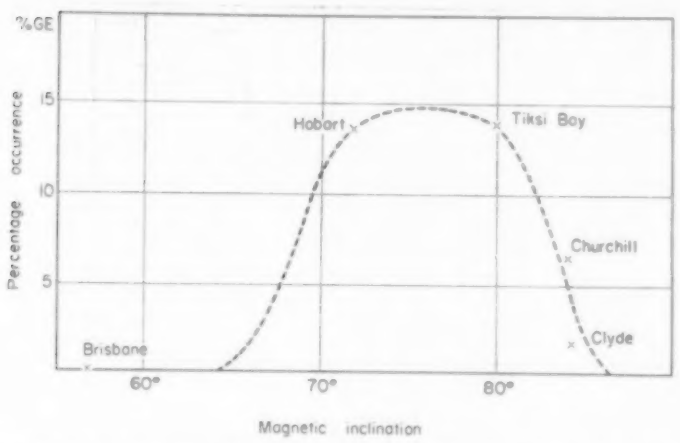


Fig. 4. Variation with magnetic inclination of the percentage occurrence of *z*-rays during 1946.

The relative occurrence of resolved spreading and unresolved spreading for this group of ionograms is comparable with that found at Hobart when all "Spread-*F*" ionograms are considered[11].

Fig. 5(a) and (b) show clean *z*-ray traces with clean *o*- and *X*-rays, while Fig. 5(c) and (d) show *z*-rays with *o*- and *X*-ray unresolved spreading and resolved spreading, respectively. Fig. 5(e) and (f) illustrate the duplicated and diffuse *z*-ray traces, respectively.

(b) *z-Rays at a mid-latitude station (Brisbane)*

(i) *Diurnal, seasonal and sunspot-cycle variations.* The diurnal variation (Fig. 6(a)) shows a gradual rise after midnight, with a maximum just before ground sunrise. The seasonal variation (Fig. 6(b)) shows a winter maximum, a small summer sub-maximum, with minima around the equinoctial periods. The winter maximum is biased towards the early part of the year. Fig. 6(c) shows the *z*-ray occurrence to vary inversely with the variation of average sunspot number.

(ii) *Nature of triple-splitting at Brisbane.* In virtually every case of *z*-ray occurrence at Brisbane the associated *o*- and *X*-rays show a spreading of the trace. A detailed analysis of the occurrence in 1944 shows that in 20 per cent of cases this spreading appears as resolved satellite traces. Again, when all "Spread-*F*" ionograms are considered, the ratio of occurrence of resolved to unresolved spreading is of the same order as that found here for *z*-ray ionograms (cf. Ref. (11)). Occasionally the *z*-ray appears diffuse (Fig. 7(a)) but this is not nearly as pronounced as that found at Hobart. Fig. 7(b) and (c) show other examples of the *z*-ray trace associated with unresolved spreading, while Fig. 7(d–f) show resolved satellite traces present at the same time as *z*-ray occurrence. The nature of the traces on Fig. 7(e) and (f) is such to suggest that the *z*-ray occurrence is related to the *F2*-layer irregularity producing the *o*- and *X*-ray satellite traces.

(c) *F2-layer ionization contours during z-ray appearances*

(i) *Hobart.* Fig. 8(a) shows a plot of contours of equal ionization density on 4 July 1946 derived from ionograms by a true height analysis by the Schmerling[15] method. In constructing these contours it is assumed, as an approximation, that all *F2*-layer reflections are vertical. Also plotted are the *z*-ray trace lengths indicated by the ionograms. (It is very seldom that the complete *z*-ray trace is present). If lines are drawn in the *F2*-layer joining the

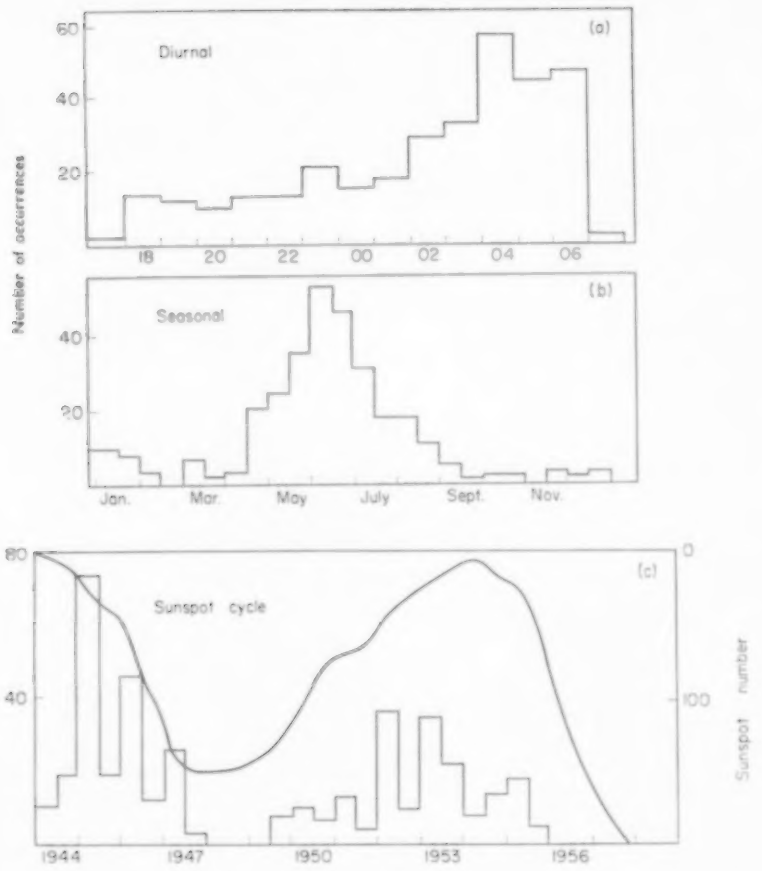


Fig. 6. Occurrence of z-rays at Brisbane 1944–1958 (inclusive). Ionograms available every 15 min 1944–July 1947; every 10 min July 1947–1958. (a) Diurnal variation. (b) Seasonal variation. (c) Sunspot-cycle variation.

position of troughs and crests, respectively, of the *F2*-layer ionization kinks, and these lines are extended to ground level, an apparent association between the upward slopes of ionization contours and the occurrence of *z*-rays, is revealed. Fig. 8(b) shows a similar plot for 11 August 1946.

The plot of *F2*-layer ionization contours in Fig. 8(c) was made because of a single occurrence of a *z*-ray trace at 17.30 on 25 July 1946 (in the sunset period). Here there is a gradual rise in the ionization contours for 30 min prior to the *z*-ray occurrence.

(ii) *Brisbane.* Ionization contours have been drawn for 28 August 1953 from 20.20 until 21.40 (Fig. 9). Resolved range-spreading traces appear for 21.42, and a *z*-ray trace appears for the same time (Fig. 7(e)). The *F2* irregularity shown by Fig. 9 is very characteristic of the "Spread-*F*" irregularities reported in Ref. (11).

3. DISCUSSION

(a) *Mechanism for return of z-ray*

It has been shown in Section 2 that evidence exists for associating the *z*-ray reflection with

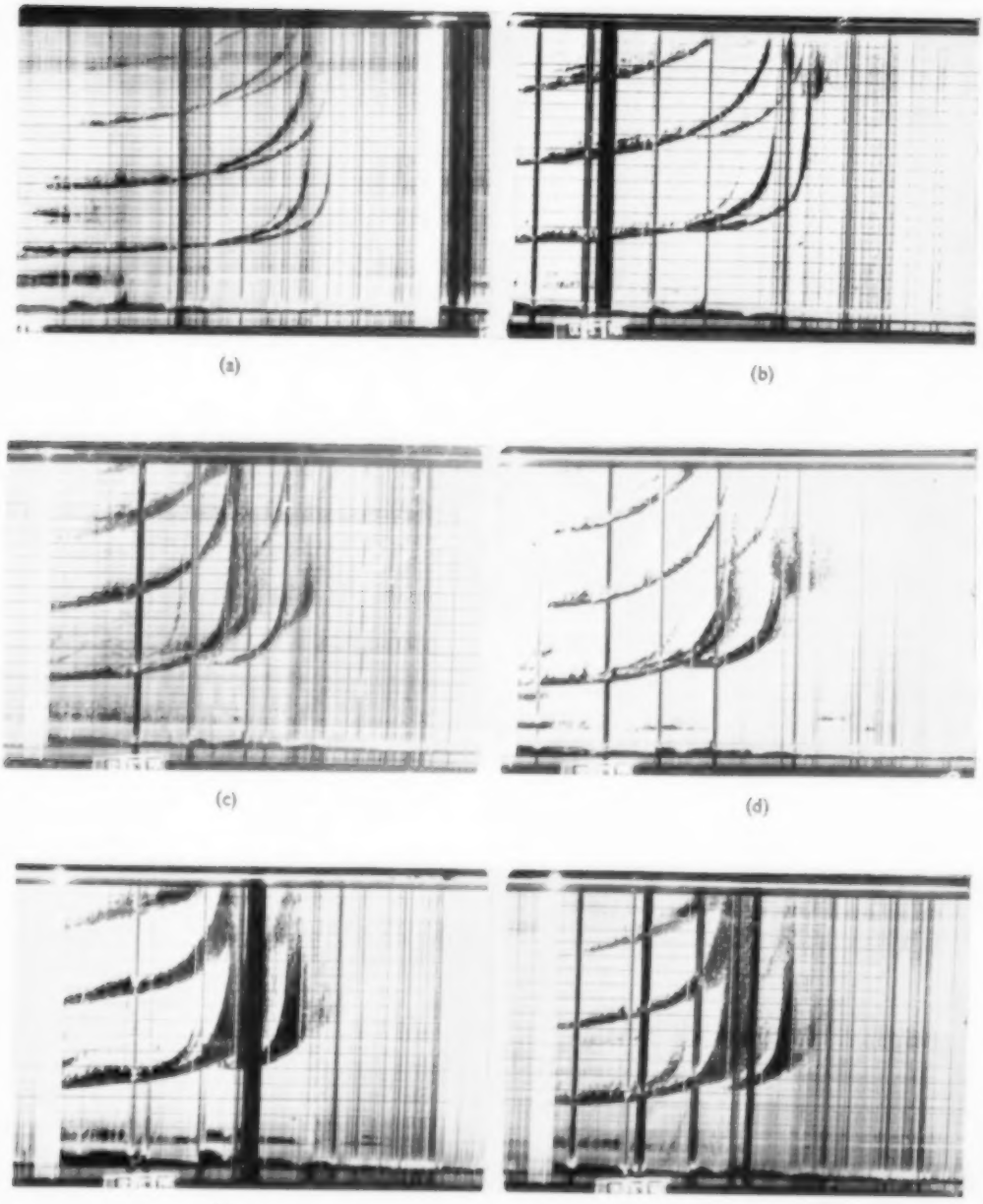


Fig. 5. Various aspects of z-ray trace occurrence at Hobart.

(facing p. 218

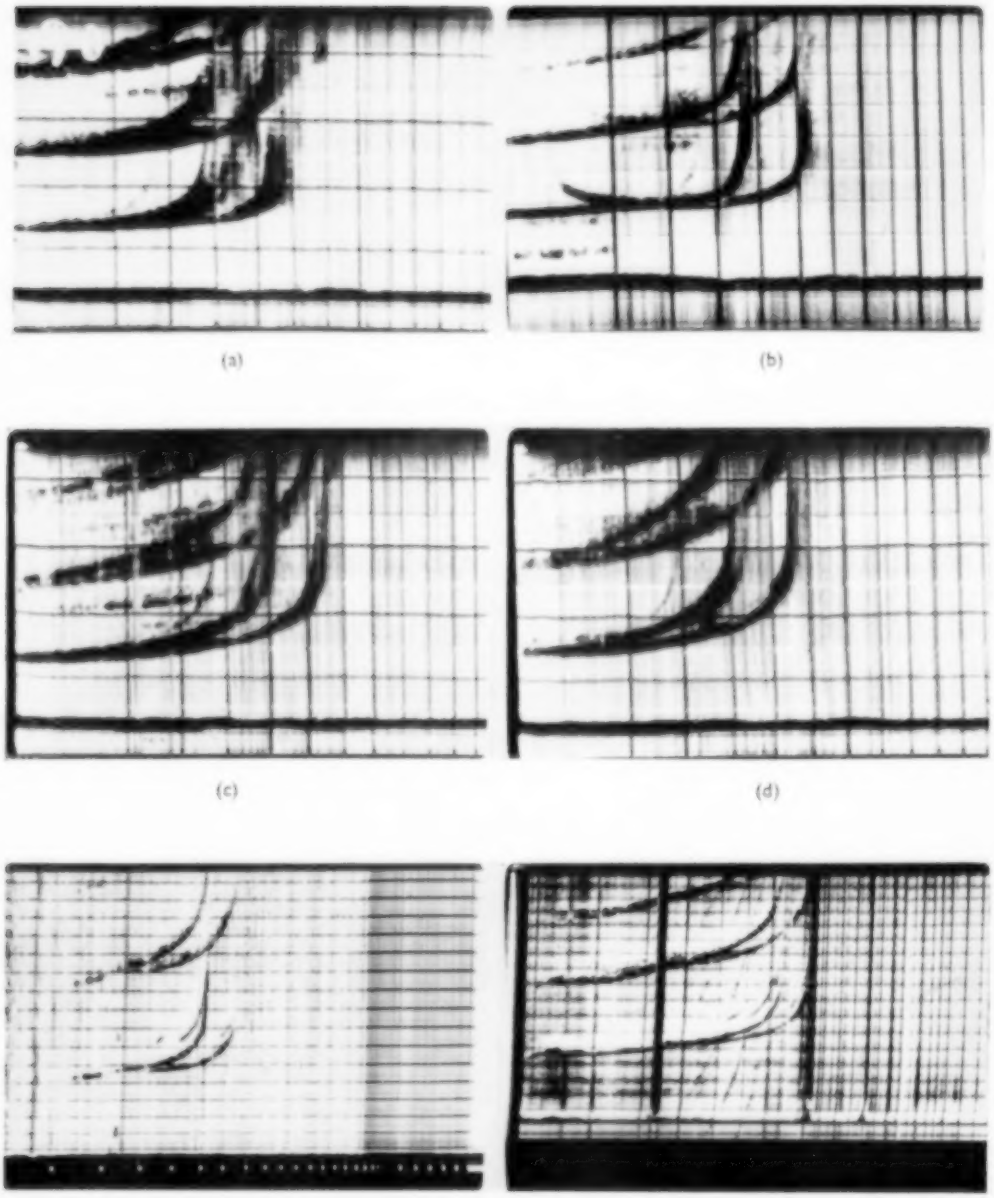


Fig. 7. Various aspects of z-ray trace occurrence at Brisbane.

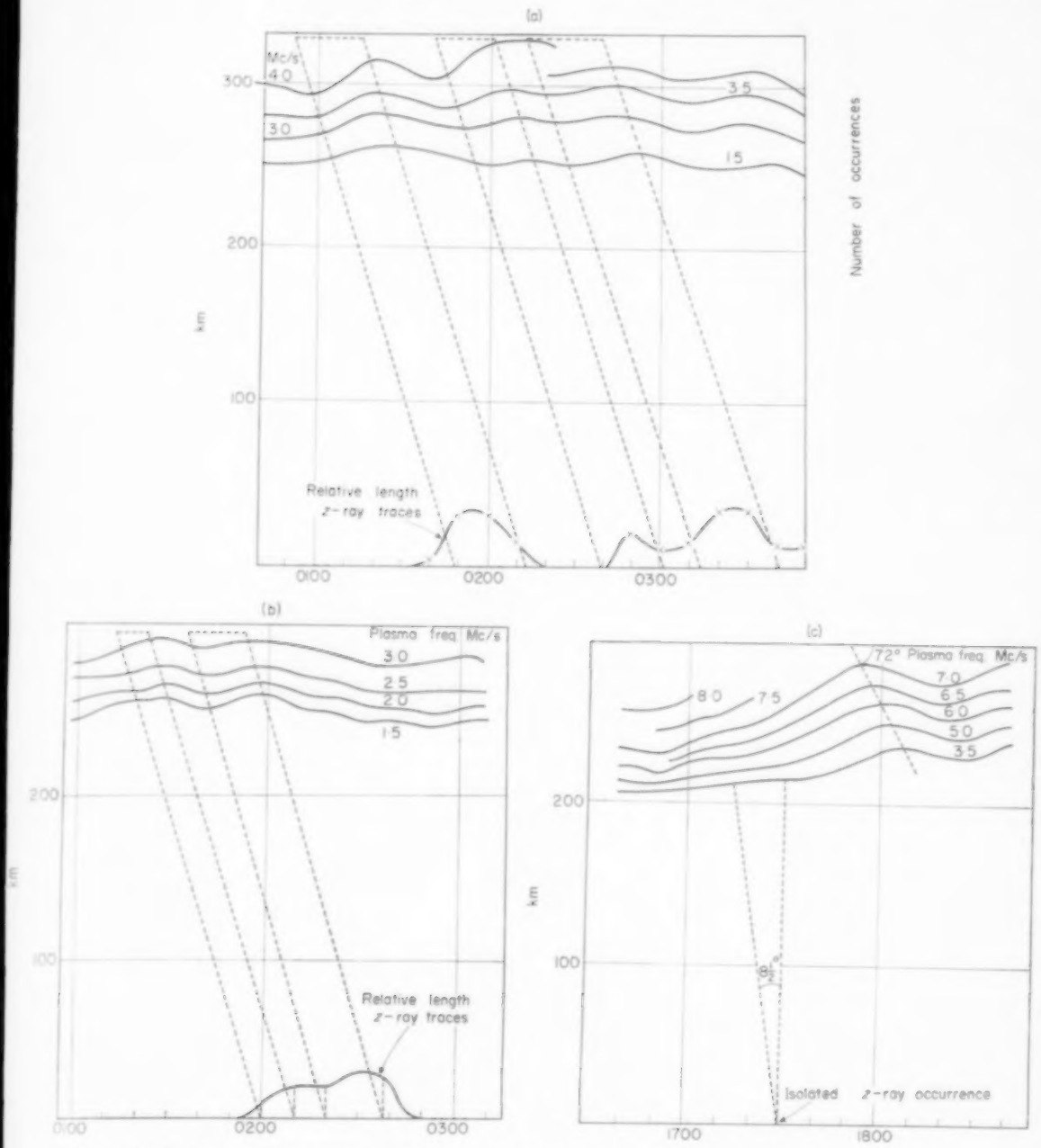


Fig. 8. (a) *F2*-layer ionization contours and z-ray trace lengths around 02.00 on 4 July 1946 at Hobart. (b) *F2*-layer ionization contours and z-ray trace lengths around 02.00 on 11 August 1946 at Hobart. (c) *F2*-layer ionization contours at Hobart around 17.30 on 25 July 1946, the time of an isolated occurrence of a z-ray trace.

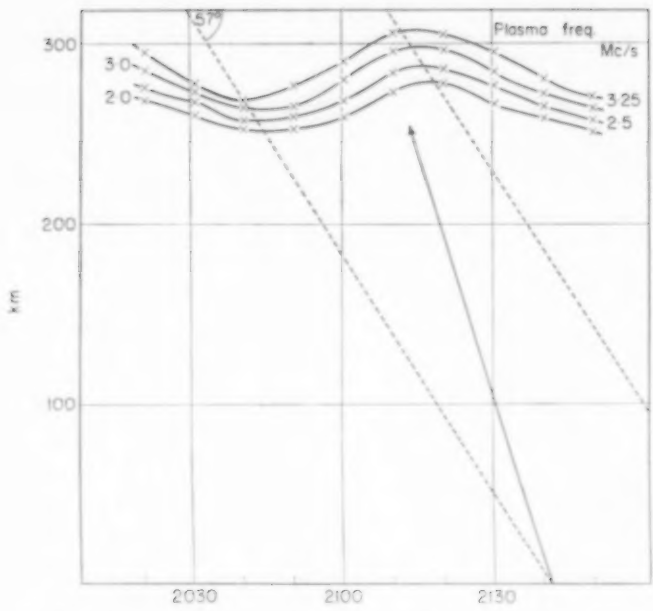


Fig. 9. *F2*-layer ionization contours apparently associated with resolved range-spreading and *z*-ray occurrence at 21.42 on 28 August 1953, at Brisbane.

the irregularities of the *F2*-layer which are responsible for "Spread-*F*" (cf. Ref. (11)). For Brisbane, the diurnal, seasonal and sunspot-cycle variations for *z*-rays are very similar to those for "Spread-*F*"[11]. For Hobart (Fig. 2) the seasonal variation is the same as that for "Spread-*F*" at this station[12]; also for Macquarie Island (Fig. 2) the winter maximum and summer minimum of *z*-rays is the same as that found at high latitudes for "Spread-*F*"[13]. Also, the same inverse sunspot-cycle relationship exists at very high latitudes (e.g. at Churchill and Tiksi Bay on Fig 3) for *z*-rays as that found for "Spread-*F*" and *z*-rays at mid-latitudes (e.g. at Brisbane on Fig. 6(c)). However, it has been shown[13] that "Spread-*F*", at very high latitudes, varies directly with sunspot activity. ("Spread-*F*" at very high latitudes may be caused by another agency, perhaps incoming streams of protons, which are directly related to sunspot activity).

These results suggest that the return of the *z*-ray results from a kinking of the ionization contours of the *F2*-layer such as is responsible for "Spread-*F*" at mid-latitudes. These irregularities are frontal in nature, extending sometimes as far as 1000 km. If the fronts are aligned perpendicularly to the magnetic meridian the *o*-wave, which has been converted to an *X*-wave by coupling in the manner described by Ellis[1], can possibly be reflected back along its path, in the plane of the magnetic meridian, because of the sloping ionization contours of the irregularity. The directions-of-arrival measured by Ellis[7], around the sunset period, suggest that the ionization contours remain approximately horizontal up to the normal *o*-ray reflection level. However, all that is required is that the ionization contours should be so shaped that a ray, which is longitudinal at the normal *o*-ray reflection level, should be normal to the ionization contours when it reaches the vicinity of the *z*-ray reflection level.

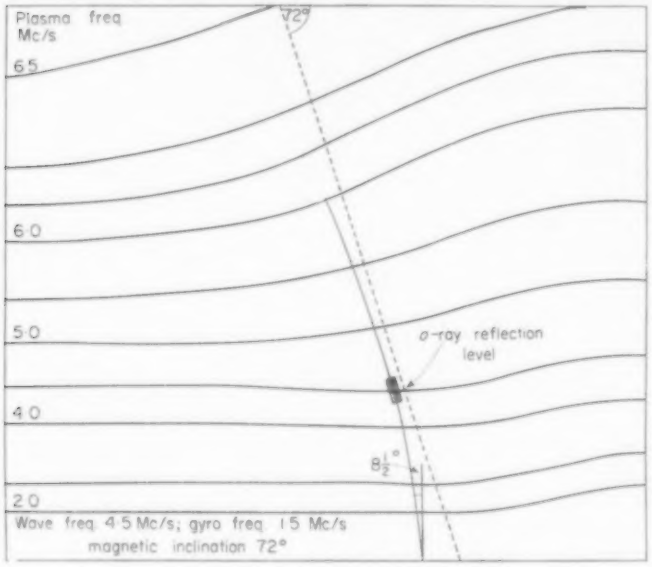


Fig. 10. Ray-path drawn in a model of ionization distribution similar to that shown in Fig. 8 (c), illustrating proposed reflection path for z-ray.

Using the refractive index diagrams, drawn in Fig. 3 and Fig. 4 of Ellis[1], ray-paths have been traced in model ionization distributions of two types. In Fig. 10 the contours are shaped somewhat similarly to those giving rise to a *z*-ray at 17.30 on 25 July 1946 (Fig. 8(c)), while in Fig. 11 the ionization contours are similar to those for "Spread-*F*" irregularities. In both cases ray-paths can be drawn giving longitudinal propagation at the normal *o*-ray reflection level and normal-incidence reflection of the *z*-ray at a higher level. These calculations have been made for a magnetic inclination of 72° and a gyro-frequency of 1·5 Mc/s (Hobart). The wave frequency assumed was 4·5 Mc/s.

In Fig. 10 the angle of incidence is the same as that measured by Ellis[7] in the sunset period, but in Fig. 11 the zenith angle of the incident radiation is somewhat larger.

The diurnal variation of *z*-rays at Hobart (Fig. 1) is not the same as that for "Spread-*F*" at this station (Fig. 27 of Ref. 11). However this is not inconsistent with our hypothesis, since only "Spread-*F*" fronts having a nearly E.–W. orientation would give *z*-ray reflection.

If these E.–W. fronts are closely spaced it is possible that two or more paths for *z*-ray reflection can exist at the same time. This would seem to explain the duplicated and diffuse *z*-ray traces reported for Hobart in Section 2.

The vertical-incidence coupling theory[3, 4] should give an increasing occurrence of *z*-rays as the magnetic pole is approached. The falling off of *z*-ray occurrence with magnetic inclination, shown in Fig. 4, suggests that this mechanism plays little part in the production of *z*-rays at very high latitudes.

(b) *Association with "Dawn Chorus"*

The diurnal variations of *z*-rays at Hobart and Macquarie Island are found to have similar distributions to those for V.L.F. emissions ("dawn chorus") at these stations; also the variations of *z*-ray occurrence and "dawn chorus" occurrence with magnetic inclination appear similar[14]. Further, the variation of the

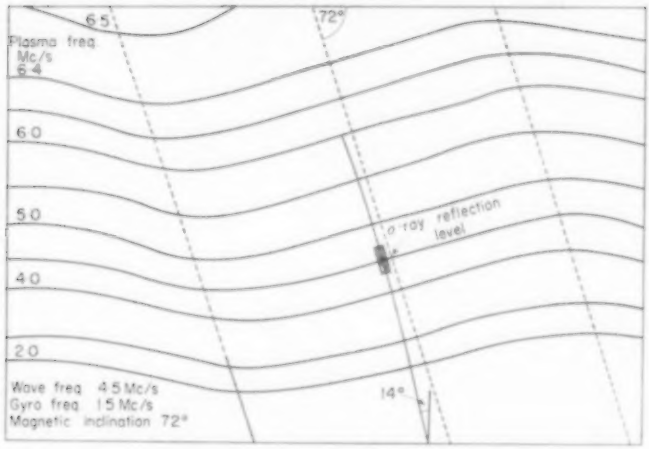


Fig. 11. Ray-path drawn in a model of ionization distribution similar to that found for "Spread-*F*" irregularities, illustrating proposed reflection path for z-ray.

time of maximum occurrence with geomagnetic latitude for z-rays is approximately the same as that found by Crouchley and Brice[14] for V.L.F. emissions.

It therefore seems likely that "Spread-*F*" irregularities, which have their frontal orientations aligned perpendicularly to the magnetic meridian, are in some way associated with the reception of V.L.F. signals from the exosphere.

Acknowledgments—The author would like to thank Professor H. C. Webster for helpful advice in the preparation of this paper. This work was financially supported by the Radio Research Board of C.S.I.R.O. and was carried out while the author was the holder of a C.S.I.R.O. Studentship. The Ionospheric Prediction Service kindly supplied the ionograms.

REFERENCES

1. G. R. Ellis, *J. Atmos. Terr. Phys.* **3**, 263 (1953).
2. R. Satyanarayana, K. Bakhru and S. R. Khastgir, *J. Atmos. Terr. Phys.* **13**, 201 (1959).
3. T. L. Eckersley, *Proc. Phys. Soc., Lond.* **B63**, 49 (1950).
4. O. E. H. Rydbeck, *J. Appl. Phys.* **21**, 1205 (1950).
5. J. C. W. Scott, *J. Geophys. Res.* **55**, 65 (1950).
6. G. R. Toshniwal, *Nature, Lond.* **135**, 471 (1935).
7. G. R. Ellis, *J. Atmos. Terr. Phys.* **8**, 43 (1956).
8. R. B. Banerji, *Indian J. Phys.* **26**, 28 (1952).
9. R. W. E. McNicol, H. C. Webster and G. G. Bowman, *Aust. J. Phys.* **9**, 247 (1956).
10. A. S. Besprozvannaya and V. A. Lovcova, *Trans. Arct. Res. Inst., N. Sea Route Auth., USSR* (1956), translated by E. R. Hope for Def. Res. Bd., Canada, July (1957).
11. G. G. Bowman, *Planet. Space Sci.* **2**, 133 (1960).
12. G. G. Bowman, M.Sc. thesis, University of Queensland (1957).
13. D. G. Singleton, *Aust. J. Phys.* **10**, 60 (1957).
14. J. C. Crouchley and N. M. Brice, personal communication.
15. E. R. Schmerling, *J. Atmos. Terr. Phys.* **12**, 8 (1958).
16. J. R. Jeffrey, *Aust. Nat. Antarct. Res. Exped. Inter. Rep.* no. 6 (1953); G. Major, *ibid.* no. 9 (1956); A. S. Little, *ibid.* no. 10 (1956).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 223–227. Printed in Great Britain.

SOLAR RADIO EVENTS AND GEOMAGNETIC STORMS

L. D. DE FEITER, A. D. FOKKER, H. P. Th. VAN LOHUIZEN and J. ROOSEN

Nera Observatory, Ionosphere and Radio Astronomy Section of the Netherlands PTT, The Hague

(*Received* 10 *February* 1960)

Abstract—On the basis of a 24-hr patrol of solar radio noise established since the beginning of the International Geophysical Year, an identification is attempted of those solar flares which, on account of their associated radio responses, most probably were the cause of a geomagnetic storm. The cases for which we think the identification to be reliable are listed. It has appeared that great integrated intensity of the radio outburst at centimeter, decimeter and meter wavelengths is the primary criterion for identifying the solar flares responsible. Most of these giant radio outbursts, to which we assigned the "radio importance" figure 3+, belong to the so-called type IV. Only a minor fraction of these events were accompanied by slow-drift bursts of type II. Of the importance 3+ radio outbursts about 60 per cent are clearly associated with the subsequent sudden commencement of a geomagnetic storm. Conversely, about 50 per cent of the sudden commencements of a storm can be related to an important radio event. Some reasons, why in a particular case the storm–outburst association may fail to exist, are mentioned.

1. INTRODUCTION

During recent years of great solar activity it has become clear that the information on solar radio outbursts associated with solar flares is of great help in identifying those solar flares which are held to be responsible for the sudden commencement of geomagnetic storms. Studies devoted to the problem how to relate geomagnetic storms to flare-associated radio events have been published, among others, by Dodson and Hedeman[1] and by Sinno and Hakura.[2]

As there are various types of solar radio outbursts, there is the interesting question whether one of these types in particular is related with the process of expulsion of the corpuscular streams which cause a geomagnetic storm. In this respect two types especially come into consideration:

(*i*) slow-drift bursts (spectral type II) with a duration of the order of 5 min,

(*ii*) strong enhancements with a continuous spectrum (type IV) and a duration of 15 min or more.

These two types sometimes occur in combination with each other, as has been pointed out several years ago by Dodson *et al.*[3]. The type II outburst constitutes the "first part" event, being followed by a strong rise of intensity, generally of type IV ("second part" event).

That a slow-drift burst might be the most relevant indication of the passage of an auroral stream through the solar corona has been advocated by Maxwell *et al.*[4]. The frequency drift of these bursts may be interpreted as corresponding with an outward motion of a disturbing agent. According to this hypothesis the velocity would be of the order of 1000 km/sec, which is just about the order of magnitude of the velocity required to travel the distance sun–earth in 1 or 2 days. The hypothesis might be tested by investigating whether there exists a systematic tendency of geomagnetic storms to occur within a few days after the occurrence of a slow-drift burst. Maxwell *et al.* found that 45 per cent of the slow-drift bursts recorded with the sweep-frequency equipment at Fort Davis (Texas) were actually associated with a subsequent geomagnetic storm.

The problem of identification of the particular solar radio event responsible for a geomagnetic storm meets two difficulties:

(1) If the observational material covers only part of the day one must allow for events

having occurred during the other part of the day.

(2) If important radio events of different types (such as type IV or II) occurred in close succession some 1 or 2 days preceding a geomagnetic storm, it is hard to decide which one is to be related with the storm.

2. CONTINUOUS PATROL OF SOLAR RADIO EMISSION

The first one of the above difficulties can be overcome by a 24-hr continuous patrol of solar radio emission. At the frequencies 200 and 545 Mc/s such a patrol has been established by our institution, the Ionosphere and Radio Astronomy Section of the Netherlands PTT (Observatory Nera), in co-operation with the observing stations at Paramaribo (Surinam) and Hollandia (Netherlands New Guinea). The latter two stations are being run by the local telecommunications services. All the equipment in use at the three stations has been developed under the supervision of F. R. Neubauer. Measurements at the two frequencies 200 and 545 Mc/s do not permit to decide whether or not a particular outburst should be classified as a slow drift burst. But they do make it possible to ascertain whether a radio event has been "important" or not.

We have supplemented our own data on 200 and 545 Mc/s with information on microwave frequencies, which is available from a world-wide chain of observatories, viz. the observatories at Ottowa, Tokyo, Toyokawa, Berlin and Nera. The complete 24-hr information has enabled us to prepare a list of all flare-associated radio outbursts that have occurred at one or more of the frequencies 200, 545 and 3000 Mc/s since the beginning of the IGY (unpublished). On the basis of the available information we have attached an importance rating to each radio event contained in the list. The category of greatest importance (which we call 3+) contains outbursts that occur with great integrated intensity at all three frequencies considered. Many of these importance 3+ outbursts belong to spectral type IV. Some of them, however, appear at metre wavelengths as a first part (possibly type II) event or as a combined (first + second) part event. In most cases the optical importance of the pertinent flares is 3 or 3+, but some of the flares are of smaller importance.

3. GREAT RADIO EVENTS AND SUDDEN COMMENCEMENT GEOMAGNETIC STORMS

Many of the great radio events (importance 3+) are followed, at time intervals of between 19 and 60 hr, by a sudden commencement geomagnetic storm. In a number of cases it is quite obvious that only one particular importance 3+ outburst comes into consideration as having been associated with the process on the sun responsible for the geomagnetic storm. In most other cases it is quite probable that the association is real. In a relatively small number of cases there is another, somewhat less important radio event which also comes into consideration for being related with a particular geomagnetic storm.

Though in individual cases one cannot always be quite certain about the reality of the association, the overall evidence strongly suggests that great solar radio events are the phenomena most closely associated with geomagnetic storm sudden commencements (SSC's). We want to stress that this is an association on the basis of the frequency extent and the total energy content at individual frequencies of the outburst.

Apparently, only a minority of these outbursts had slow-drift characteristics. There are methods to verify the significance of the association statistically (De Feiter and Roosen[5]). Their application, however, would lead us into too many details in the present context. We just summarize our findings, which pertain to "*bona fide*" SSC-storms. As such were considered storms with maximum magnetic index K_p greater than or equal to 5+, that were observed by a sufficiently great number of observatories. We adopted 60 hr as the greatest admissable delay between flare and SSC.

(1) Among the fifty radio importance 3+ events thirty-two, or 64 per cent, can be

Table I. Proposed Associations of Geomagnetic Storms with Solar Flares

Storm sudden commencement				Flare				ΔT (hr)	Affirmation
Date	Time (U.T.)	K_p^{max}	$(\Sigma K_p)_{max}$	Date	Time	Position	Imp.		Reid and Leinbach[6] (o) Hakura and Goh[7] (x) Maxwell *et al.*[4] +
1957									
Apr. 17	11·36	8—	41	Apr. 16	10·40	31N 85E	3	25	
Apr. 18	15·08	7—	44+	Apr. 17	20·00	27N 70E	3	19	+
Jun. 30	05·28	8+	61+	Jun. 28	06·58	9N 27E	2	46	
				Jun. 28	12·23	12N 21E	2	41	
Jul. 4	23·42	7+	38—	Jul. 3	08·30	9N 42W	3	40	o x
Aug. 29	19·09	7—	40+	Aug. 28	09·13	30S 35E	3+	34	x
Aug. 31	18·12	7	40+	Aug. 30	22·15	SID	3	20	+
Sep. 2	03·14	9—	59—	Aug. 31	12·57	25N 2W	3	38	x
Sep. 4	13·00	9	62+	Sep. 2	12·57	11N 26W	2	48	o x
Sep. 12	21·54	9—	54+	Sep. 11	02·43	11N 3W	3	43	x
Sep. 22	13·45	9—	60+	Sep. 21	13·30	11N 7W	3	24	o
Sep. 29	00·16	9—	56	Sep. 26	19·07	25N 15E	3	53	o
Oct. 21	22·41	7—	32—	Oct. 20	16·37	27S 38W	3+	30	o x
Nov. 6	18·21	7	36+	Nov. 5	12·05	24S 53W	2	30	
Nov. 26	01·55	7—	46—	Nov. 24	08·48	14S 36E	3	41	x
Dec. 31	05·14	6+	45—	Dec. 28	22·29	25N 50W	2	55	+
1958									
Feb. 11	01·25	9	60—	Feb. 9	21·08	12S 14W	2+	28	o x
Mar. 5	05·37	7—	38	Mar. 3	10·05	17S 60E	3	43	
Mar. 25	15·40	6	32—	Mar. 23	09·47	14S 80E	3+	54	o x
				Jun. 4	21·47	15N 58W	2	51	x
Jun. 7	00·46	8—	44+	Jun. 5	16·56	21S 67E	2+	32	x
				Jun. 6	04·36	16N 78W	2	20	x
Jul. 8	07·48	9	62—	Jul. 7	00·20	25N 8W	3+	31	o x
Jul. 21	16·37	7+	39	Jul. 19	19·05	23N 13E	2+	45	+
Aug. 17	06·22	7+	48	Aug. 16	04·32	14S 50W	3+	26	o x
Aug. 22	02·27	6—	31+	Aug. 19	21·18	18N 26E	2	53	
Aug. 24	01·40	8—	44+	Aug. 22	14·17	19N 9W	3	35	o x
Aug. 27	03·03	7—	42+	Aug. 26	00·05	20N 54W	3	26	o x
Sep. 16	09·30	5+	38—	Sep. 14	08·22	10S 80W	2	49	x
Oct. 24	07·30	7	49—	Oct. 21	23·18	4S 22W	2	56	x
Dec. 13	00·01	6	39	Dec. 11	18·02	2S 0	2	30	x
1959									
Feb. 14	11·42	6—	36	Feb. 12	23·01	13N 48E	3	37	
Apr. 9	18·29	8—	49	Apr. 8	09·05	27N 85E	3	33	
May 11	23·26	8+	52+	May 10	20·55	20N 48E	3+	26	o
Jun. 11	09·09	6	23	Jun. 9	04·45	SID	3+	52	
Jul. 11	16·26	7—	37	Jul. 10	02·10	20N 66E	3+	38	o
Jul. 15	08·03	9	64+	Jul. 14	03·25	16N 5E	3+	28	o
Jul. 17	16·39	9—	59+	Jul. 16	21·14	15N 30W	3+	19	o
Aug. 16	04·04	8+	36+	Aug. 14	00·40	10N 28E	2+	49	
Aug. 20	04·12	6—	36+	Aug. 18	10·13	13N 32W	3	42	

related to an SSC-storm with a high degree of probability.

(2) Of the seventy-six *"bona fide"* SSC-storms thirty-seven, or 49 per cent, were the obvious consequence of a solar process that gave rise to an important radio event.

(3) The mean time interval between the outburst and the associated SSC is 37 hr.

These results pertain to the period April 1957–December 1959.

4. PROPOSED IDENTIFICATION OF SOLAR FLARES RESPONSIBLE FOR SUDDEN COMMENCEMENT GEOMAGNETIC STORMS

In Table 1 we have listed those solar flares which, in our opinion, can be considered as responsible for a subsequent geomagnetic storm with a high degree of reliability. A number of these flares have also been mentioned by other authors (for instance Reid and Leinbach[6], and Hakura and Goh[7]). About *these* flares there evidently exists a *communis opinio* and one may be very confident that their identification is correct. Of some of the other identifications we may have been in error. Still we think that their number is quite small.

Comparing our data with the list of slow-drift bursts tabulated by Maxwell *et al.*[4] there are five cases of which we do not think that the cited events were responsible for the subsequent geomagnetic storm. These pertain to the sudden commencements on the following dates: — 1957: 17 Apr., 29 Aug., 4 Sep. 1958: 11 Feb., 17 Aug. The flares in Table 1 caused radio events which were more important than the proposed slow-drift bursts.

We do agree as to the sudden commencements of 1957, 18 April, 31 August, 31 December and 1958, 21 July. The other sudden commencements of their list could not be related with radio events of great importance.

It seems that a proviso should be made to the delay of the magnetic storm relative to the outburst. Judging from the minimum time delay (19 hr), found for the well-established SSC-outburst associations, we doubt if any identification implying time delays of say, less than 15 hr, is correct. For instance, we do not think that the flare of 1957, 12 September, 15·6 U.T. caused the SSC of 12 September, 21·54 U.T., though the associated radio response was very great indeed.

5. DISCUSSION

The fact that many sudden commencement geomagnetic storms correspond with great radio outbursts (generally belonging to type IV) seems to indicate that radio events of the greatest importance class are excited by the very corpuscular streams that cause geomagnetic disturbances on the earth. However, in almost 40 per cent of all cases an importance 3+ radio outburst is *not* followed within 60 hr by an SSC of some importance. One of the reasons, why in a particular case the pertinent flare fails to cause a geomagnetic storm, may be connected with the well-known effect of seasonal variation of geomagnetic activity. Obviously, the conditions for the occurrence of geomagnetic storms are less favourable in the periods around the summer and winter solstices than around the equinoxes. In this connection we mention the following great radio events for which association with an SSC storm might normally have been expected:

(*a*) Important radio flares which failed to cause a geomagnetic storm:
11 May 1959, 20·06; 17 May 1959, 05·23; 15 January 1960, 13·35.

(*b*) Important radio flares followed by a minor magnetic storm which did not meet our requirements as to strength and/or time delay:
16 July 1957, 17·42; 24 July 1957, 17·12; 14 December 1957, 12·45; 17 December 1957, 07·34; 6 June 1958, 04·36; 23 December 1958, 05·45.

Conversely, about one half of the geomagnetic storms cannot be related to a particular radio outburst of great magnitude. Still sometimes there has occurred an important optical flare which might be considered as responsible. In this connection we must allow for the possibility that a second part flare-associated radio event is inaccessible to observation.

There are reasons to believe that enhanced radiation at metre wavelengths may be screened by coronal structural elements, especially if the source is situated near the limb (Fokker[8]). Under certain circumstances prolonged radio emission associated with corpuscular active flares therefore might remain unobserved.

Every identification of the solar process responsible for a geomagnetic storm is to some extent tentative. Due to the multitude and complexity of solar phenomena one must be specially careful in drawing conclusions. In this connection we want to emphasize one point in particular.

The different solar phenomena all have a tendency to occur during the same periods, when solar activity is on a high level. During such periods there is a fair chance that the solar phenomenon X, which brings on a geomagnetic storm, will occur. In this case there is also an enhanced probability of occurrence of other primary phenomena, A, B, etc. Concentrating now on events of type A, we may find a quasi-correlation with magnetic activity. For instance, using the method of superposed epochs, one might find a spurious effect. An increase of the average magnetic index K_p, some 2 days after the occurrence of events of type A, does not prove that *these* events are actually the ones responsible for magnetic disturbances. The only relevant procedure is to compare several superposed epoch diagrams pertaining to different categories of events.

REFERENCES

1. H. W. Dodson and E. R. Hedeman, *J. Geophys. Res.* **63,** 77 (1958).
2. K. Sinno and Y. Hakura, *Rep. Ionosphere Res. Japan* **12,** 285 and 296 (1958).
3. H. W. Dodson, E. R. Hedeman and L. Owren, *Astrophys. J.* **118,** 169 (1953).
4. A. Maxwell, A. R. Thompson and G. Garmire, *Planet. Space Sci.* **1,** 325 (1959).
5. L. D. de Feiter and J. Roosen, in preparation.
6. G. C. Reid and H. Leinbach, *J. Geophys. Res.* **64,** 1801 (1959).
7. Y. Hakura and T. Goh, *J. Radio Res. Lab. Japan* **6,** 635 (1959).
8. A. D. Fokker, in preparation.

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 228–237. Printed in Great Britain.

ARTIFICIAL ELECTRON CLOUDS—VI

LOW ALTITUDE STUDY, RELEASE OF CESIUM AT 69, 82 AND 91 km

J. PRESSMAN,* F. F. MARMO,* and L. M. ASCHENBRAND*

Geophysics Corporation of America, Boston 15, Massachusetts

(*Received* 23 *July* 1959)

Abstract—There is reported a series of three artificial electron clouds created at Holloman Air Force Base, New Mexico at 69, 82 and 91 km in May 1958 by the day-time thermo-chemical release of cesium. The electron cloud was created by both thermal and photo-ionization. A lower limit for the effective duration of electron clouds for this kind of release was found to be 70 km. The short duration at low altitudes is caused by both the rapid chemical consumption of the cesium atoms thereby preventing photo-ionization and the rapid electron attachment followed by mutual neutralization. The coefficient of mutual neutralization is estimated to be 10^{-7}–10^{-8} cm^3 sec^{-1}. The mathematical model employed appears to be reasonably valid.

1. INTRODUCTION

A previous paper in this series[1] discussed the general considerations involved in the physics of low altitude (below 90–100 km) artificial electron clouds and in particular the generative and decay processes associated with the history of such representative electron clouds. Several cases were presented based on these theoretical considerations, and a simple analysis was made of those processes which were more dominant (e.g. chemical consumption, recombination, electron attachment, etc.), at the low level in contra-distinction to those other processes which were more important above the 100-km regime.

This paper reports on a series of three experiments that were performed for the purpose of studying the effect of low (below 100 km) altitude on the formation and history of artificial electron clouds. These flights were designed to investigate not only the effect of individual processes but also the earlier conclusion[1] that there probably exists some lower altitude below which the generation of a detectable, high density, long-lived, solar-induced artificial electron cloud could not be achieved with the experimental techniques employed here. The designed approach was as follows. The first experiment was to be conducted at 80 km and if positive radar detection results were obtained, a similar release at 70 km would be attempted. If the latter results were positive, then a release at 60 km would be performed. However, if the results of the 70 km release proved to be negative, a release at 90 km would be attempted in order to complete this investigation.

Additionally, a model of a reasonably close degree of applicability has been set up and analyzed mathematically so that the previous exploratory analysis in effect, could be moved forward to some extent for comparison with the experimental data where possible.

2. EXPERIMENTAL FACTS

2.1. *Launchings*

Three Nike-Cajun rocket experiments were conducted in September 1958, at Holloman AFB, New Mexico, in which cesium atoms were released into the upper atmosphere. Details pertaining to the chemistry associated with the alkali metal nitrate–aluminum mixture[2], the engineering techniques[2], the canister design[3], vehicle capabilities, etc., have been previously presented in detail in other

* Formerly with Geophysics Research Directorate, Air Force Cambridge Research Center, L. G. Hanscom Field, Bedford, Massachusetts.

This work has been partially supported under contracts: U.S. Army Signal Corps DA36-039-SC-78971 and U.S. Army Signal Corps DA36-039-SC-78925.

Original work was performed when authors were employed at GRD, AFCRC.

Table 1. Summary of Experimental Facts Pertinent to September 1958 series of Nike-Cajun Experiments

Altitude of release (km)	Date (1958)	Time of release (MST)	Release time after launch (sec)	Total chemical weight (kg)	g-atoms cesium
69	16 Sept.	0931:59	59·6	9·82	27·7
82	15 Sept.	0933:06	65·8	9·39	27·0
91	18 Sept.	0951:14	76·0	10·60	30·4

papers of this series. Table 1 presents a summary of the basic experimental facts concerning the three launchings.

2.2. *Optical detection*

The three releases were conducted in the daytime since no attempt was made to accomplish morning twilight experiments. In a previous paper[4] in which optical and radio-radar data were correlated, a discussion pertinent to the decided advantage of a visible cloud has been presented. Briefly, one can ascertain through optical and photographic techniques the precise altitude of release, winds and shear effects, cloud dynamics, chemical yield, etc. It was deemed important in these previous experiments to "dope" the clouds with sodium and to use the previously described optical detection techniques with a twilight release. However, for the present case in which relatively low altitudes of releases were needed, the existing sky background at the proper release time (at about earth sunrise) was too intense to render visible any cloud by the twilight resonance effect for any appreciable time. More sophisticated optical instrumentation may have helped, but optical detection in any event was not considered of paramount importance in this phase of the program.

In this particular series, it was especially important just to ascertain that a successful release had been accomplished at a certain altitude for the case of possible negative radar results. Consequently, another method of rendering detectable a release was attempted. This method consisted of generating a magnesium oxide smoke bomb. In particular, a mixture of 360 gr of barium dioxide and 39 gr of Mg powder was set for a squib-initiated release about 3 sec prior to the two lower (69 and 82 km) altitude cesium releases. It was considered that the small amounts of chemical used for this supplemental release would not affect the major release or subsequent photoelectron generation. In an attempt to photographically record the smoke bombs, a triangulation camera net of three 35 mm Beattie–Coleman Varitron[3] cameras and Press Cameras[3] was utilized. The cameras were equipped with adequate filters and fast film but the effect was not photographically recorded. However, visual (and radar) observations were obtained in each case.

The inability to photograph the MgO puff could be ascribed to the utilization of an insufficient amount of material. In addition, the material was probably dispersed throughout too large a volume, thus so severely attenuating the MgO density that it was not photographable against the daylight sky background. Successful optical coverage could probably be obtained by utilization of more material or utilizing more sophisticated optics.

Although it appears that the scheme for generating a visible MgO puff failed optically, it was a fortunate fact that it was successful in another, quite unexpected manner. This was involved in detection of the generated pilot cloud by the Stanford University radar sytem at Alamagordo, New Mexico. This is discussed further in the next section.

2.3. *Radio-radar detection*

A description of much of the radar ground detection devices has been given elsewhere[1, 2, 3, 4]. The following units were located 15 miles east from launch site: (a) a 40·7 Mc, 50 kW, pulsed radar with a minimum

detectable cross-section capability of 4 m², (b) a 20·8 Mc, 2·8 kW, pulsed radar with minimum detectable cross-section capability of 215 m², (c) at 6 Mc, 21 kW, pulsed radar with a 120 m² cross-section capability (d) a 100 kc, 1·8 kW pulsed probe with a minimum cross-section capability of 1 km², and a low sensitivity to small cross-sections. No results were obtained on this latter instrument. Additionally, the *C*-3 ionosonde (1–25 variable frequency sweep), located at White Sands Proving Ground, New Mexico, obtained good data. The Laredo and Ft. Huachuca nets and related forward-scatter experiments were not utilized since the low release altitudes employed place the electron clouds below the antenna patterns of the equipment at these installations. For all three experiments, a 23·1 Mc, 70 kW, pulsed back-scatter radar was operated by Stanford University near Tucson, Arizona. An additional ground radar system which has not been previously described was used in this series. This system consisted of two portable Ionospheric Recorders of Cossor Ltd., which are capable of recording automatically, at predetermined intervals, panoramic records of ionospheric soundings over a variable frequency band of 1·6 to 20 Mc/sec. The results are recorded on 35 mm film together with the time and data of the event. These recorders were located 32 miles east of the launch site.

3. RESULTS

On September 15 a solar flare occurred which contributed an additional effect to the creation of the artificial cloud. A check on the occurrence of this flare[5] established the time as at least 5 min after the chemical constituent release at altitude. It will be shown by the positive radar responses associated with the electron cloud creation that the instrumentation recorded the natural solar effects only after about 10 min following the electron cloud formation. The evaluation and subsequent treatment of the data for this date (82 km release) is therefore considered reliable only for the first 10 min after release.

The radar results are presented in a format which concisely describes the reliable data for each of the experiments per detection unit. It shall be noted that the report of a negative result on the *C*-3 ionospheric units signifies that no detection was obtained in any (especially the first) of 15-sec frequency sweeps. This does not rule out the possibility of an effective duration of less than 15 sec since these instruments might not necessarily record short durations.

3.1. *Cossor portable C-3 ionosonde recorders*

(*a*) 69 km—no results were associated with the cloud formation.

(*b*) 82 km—an echo was recorded 10 min after release at a slant range of about 100 km. This echo persisted for about 20 min with a maximum frequency response of about 6 Mc appearing at *x* plus 18 min. The nature of the results (time of event and duration) are contrary to both theory of cloud formation at this altitude and results from other units. This data is considered to be probably due to the solar activity and aids in establishing the time of effect of the flare.

(*c*) 91 km—results were obtained for a definite duration of 2 min for 6 Mc and a probable duration of 7 min for over 4 Mc.

3.2. *White Sands C-3 ionospheric station*

(*a*) 69 km—no results of cloud detection were obtained.

(*b*) 82 km—an echo at a slant range of about 118 km was immediately detected following constituent release at a maximum frequency of 6 Mc. Thereafter, the echo increased in maximum frequency up to 9 Mc after 4 min. Approximately 5 min later, the effect of the solar flare was detected, therefore, only the pre-flare results were considered reliable.

(*c*) 91 km—an intense echo was recorded during the first 15-sec sweep after release at an initial slant range of 120 km. Frequencies of up to 20 Mc were recorded, and the echoes persisted for about 5 min. Thereafter, the results were considered inconclusive.

3.3. *The* 20·8 Mc *radar unit*

No signal which could be associated with the electron clouds was indicated for any of three flights.

3.4. *The* 6 Mc *radar unit*

(*a*) 69 km—no result.

(*b*) 82 km—an echo was observed at about 8 min after release and lasted about 23 min. This result is again ascribed to solar activity.

(*c*) 91 km—no results.

3.5. *The* 100 kc *radar unit*

No echoes were detected on any test.

3.6. *The* 23·1 Mc *radar located at Tucson, Arizona*

This instrument was only employed on the 91-km release experiment and a solid response lasting for about 2 min was recorded.

3.7. *The* 40·7 Mc *radar unit*

The 40·7 Mc radar data were the most consistent and compatible to analysis. These were

(*a*) 69 km—two echoes associated with the release and subsequent formation of the electron cloud were recorded. Immediately following the MgO puff release, an echo was observed for 0·6 sec at a slant range of 67 km at about plus 4 dB. 3 sec later, an echo associated with prime release was detected at a slant range of 71 km for 1·5 sec and a strength of plus 20 dB.

(*b*) 82 km—a record of an echo of 0·5 sec duration at a slant range of 78 km at about plus 5 dB strength was obtained. About 6 sec later, another response was reported to have lasted for an additional 4·3 sec at a slant range of 84 km at plus 30 dB. The latter was electron cloud generated.

(*c*) 91 km—this radar unit recorded a response at the time of contaminant release (no MgO marker employed) that was the strongest of this series. The duration of the initial signal was for 35 sec with a 25 dB strength. After this initial signal lasting for 30 sec had decayed slightly, a secondary buildup occurred about 15 sec later that lasted for about 20 sec having the same amplitude and signal strength. Thus, the total radar response for 40·7 Mc for this release was about 70 sec.

3.8. *Observations regarding overall results*

Some comments should be made about the general nature of these results. First, the 40·7 Mc echo was observed on all three experiments. The duration of the echo attributed to the cesium is consistent with the expectations (increasing strongly with altitude). The cross-section requirements of the cloud for this frequency are small, being of the order of 2–6 m^2, so that if critical electron density of $2 \times 10^7/cm^3$ is achieved, and this seems likely, echoes should be received. At 20·8 Mc the critical back-scatter electron density of $5{\cdot}4 \times 10^6/cm^3$ and a cross-section of 100–300 m^2 must be achieved. The lack of any reliable echo on this frequency is attributed to the small initial size of the electron cloud. Further, its behavior could be such that, by the time it has attained the critical size through diffusion, the electron density has dropped below the critical density. The 23·1 Mc Tucson radar return for the 91 km flight indicates the borderline nature of the physical conditions for detectability by the various pieces of apparatus used for this flight.

With regard to the Cossor and *C*-3 data, the 69 km release showing no return is as expected and consistent with the other returns for this flight. However, a comparison for the 82 km and 91 km for these instruments presents a more complicated picture. First, the longer lasting Cossor record is attributed to a strong drift of the cloud to the Cossor locale and out of the White Sands pattern. Study of the range variations of the various records is in support of this thesis. Second, while a longer enduring cloud was expected for 91 km than for 82 km and is borne out by the general nature of the radar returns, the Cossor and *C*-3 returns were less enduring for the higher altitude. This reversal is not understood at the present and will bear further consideration.

These peculiar returns may have been due to changes in the sensitivity of these recorders for which no adequate calibration exists.

In the following section a mathematical model will be set up for comparison with the foregoing results. It was felt that the best opportunity for comparison lay with the data recorded by a single instrument since these would be, at a minimum, internally consistent.

Furthermore, what was wanted was a set of positive responses for each of the flights. Also it was important that the data selected be for a reasonably high frequency so that the conditions for the cross-section be easily met, i.e. $2r_0 \geqslant 2\lambda$. If these data while being positive in addition were of short duration this would also serve to mitigate the extraneous wind effect.

The only set of data meeting these requirements was that for the 40·7 Mc radar. Hence, the theoretical analysis will be applied to these data as the most reliable and valid.

4. THEORY AND ANALYSIS

The governing processes considered here for lower altitude generation of artificial electron clouds include (1) photoionization and photodetachment, (2) electron attachment to O_2 and recombination, including mutual neutralization, (3) chemical consumption, (4) initial expansion, chemical yield and thermal ionization, and (5) diffusion both neutral and ambi-polar. These are discussed below individually.

A further consideration was the variation of these processes at the different altitudes and their relative roles in the generation and destruction of electrons. Then a suitable mathematical model was derived and applied to the data.

4.1. *Photoionization and photodetachment*

In the region 69–92 km, the probability of photoionization of Cs, k, remains relatively constant since there is little absorption by the intervening atmosphere in the wavelength region between 2000–3200 Å. At least 99 per cent of the radiation in this region is transmitted down to 40 km.[6] The k value of $6{\cdot}5 \times 10^{-4}$ sec^{-1} was used here as in previous papers of this series.[1] Similarly, the probability of photodetachment is not affected by absorption and the previous value of 0·40 sec^{-1} was used.[1] Scattering is inconsequential for both processes.

4.2. *Electron attachment and recombination*

The values used for electron attachment were taken from Fig. 4 of Paper II[1] of this series. The figure indicates that for the two-body and three-body rates of $1{\cdot}3 \times 10^{-15}$ cm^3/sec and 5×10^{-30} cm^6/sec, respectively, the latter is the dominant process below about 90 km. Thus this is the only attachment mechanism considered here. The cesium-ion recombination, α_r, was taken to be as high as 10^{-10}cm^3/sec. However, even with this value the conclusions arrived at below are valid. The coefficient of mutual neutralization (α_i) of the Cs^+ and the negative ions (O_2^-) was arbitrarily estimated at 10^{-7} cm^3/sec, since little information exists for this specific process. Values of $\alpha_i = 10^{-7}$ cm^3/sec have been considered by Bates[7] for similar processes. Order-of-magnitude values for α_i have been obtained in the analysis below.

Because of the rapidity of the electron–negative ion equilibrium under daylight conditions[1] it is valid to consider the apparent recombination coefficient[8]

$$\alpha = \alpha_r + \lambda\alpha_i \qquad (1)$$

where λ = ratio of negative ions to electrons at equilibrium.

4.3. *Chemical consumption*

The analysis of chemical consumption has been made previously.[1] For this lower altitude study the pertinent reaction of the set of reactions evaluated is

$$X + O_2 + M \longrightarrow XO_2 + M$$

We have used the consumption life-times (t_3) given in Fig. 6 of Paper II of this series where t_3 is approximately 10^{-2} sec at 60 km and increases markedly with altitude so that at 90 km $t_3 = 10^2$ seconds. It should be noted that the alkali consumption reaction rate used was that for Na while those for Cs is the pertinent one here. It is further noted that the reaction rate for Na has only been crudely measured.[1] Thus, it is felt that these adopted rates represent at best a good approximation. Finally, it is implicit in the calculations below that the concentration of Cs is less than that of the ambient molecular oxygen so that at all times there exists an adequate reservoir for consumption.

4.4. *Estimates for Gaussian half-width and initial electron density*

For these low altitude daytime experiments, the contaminant cloud was not made visible via twilight release as discussed earlier. The lack of optical data to give size and shape of the cloud made it necessary to estimate the value for the initial size or Gaussian half-width to apply to the mathematical formulation. Because of the complexity of this type of high altitude explosion, no attempt was made to solve this problem, *ab initio*. Rather the technique used, admittedly crude, was to extrapolate the previously optically measured initial sizes and scale them appropriately for low altitudes as follows. Data was available on the Gaussian half-widths, r_0, obtained for similar releases at 128[4], 116[4] and 101[3] km. Further, the latter measurements indicated that a high degree of mixing and dilution had occurred during the initial expansion as a result of the total process (i.e. turbulence, shock, instabilities, etc.) which occurred prior to the normal diffusion process. Also it was felt that, fundamentally, the initial expansion was stabilized by the engulfing of a specific mass of air and the sharing of the energy of the high temperature gases evolved. Thus, lower altitude r_0 estimates were obtained by extrapolation of the observed higher altitude r_0 values via the atmospheric density behavior from 130 km to 70 km. The best fit data yielded r_0 estimates of 117, 69 and 38 meters for the 91, 82 and 69 km, respectively.

This method contains the implicit assumption that each of the lower altitude payloads behaved identically amongst themselves and with the upper altitude releases at least in regard to the intensity of explosion, chemical yield and degree of thermal ionization. In particular, possible lack of reproducibility in the explosive release remains a factor that cannot be evaluated here. Hence, the estimated Gaussian half-widths and initial electron density values are considered to be order-of-magnitude.

4.5. *Effect of diffusion*

The general nature of the following considerations evaluate the effect of diffusion on the electron cloud, in particular the center-point electron density so as to compare it with other operative processes. It should be remarked here that there are two kinds of diffusion involved, that of the charged particles (ambipolar diffusion) and that of the neutral particles with coefficients D_a and D_n, respectively, where D_n is substantially larger than D_a.

Theoretically computed[9] values for D_a and D_n for 91 km indicate that the diffusion $t_{\frac{1}{2}}$ (the time for the center-point density to decrease by one-half) values for the charged and neutral particles are $1{\cdot}1\times10^3$ sec and $1{\cdot}2\times10^2$ sec, respectively. In addition, the respective values become increasingly much higher for 82 km and 69 km. These times should be compared to the total time of the three 40·7 Mc returns, namely, 1·5 sec, 4·3 sec, and 70 sec for the 69 km, 82 km and 91 km releases, respectively.

Thus, it is clear that the case of ambipolar diffusion need not be included in the mathematical model. However, only for the 91-km release the neutral diffusion time scale is of the order of the time scale of the phenomenon encountered in this study. Since the neutral diffusion effects the electron density by decreasing the photoionization production term, this should properly be included in the mathematical analysis for the 91-km release. In the treatment below this effect is taken into account by the introduction of an exponential term.

4.6. *Mathematical considerations*

Previous considerations[1] indicate that the time decrease of the neutral cesium density due to chemical consumption can be expressed as

$$n=n_0e^{-bt} \tag{3}$$

where n = concentration of Cs atoms, cm^{-3}
n_0 = concentration of Cs atoms at $t=0$, cm^{-3}
b = chemical consumption rate for Cs, sec^{-1}.

The mathematical formalism includes the effect of diffusion on the neutral by writing

$$n=n_0e^{-(b+\varepsilon)t}=n_0e^{-b't} \tag{4}$$

Here ε represents the relaxation time for diffusion and at 91 km is equal to 0·007 sec^{-1}.

At the two lower altitude releases ε is insignificant compared to b. Admittedly, this is only an approximtae technique for estimating the diffusive second-order effect. In the numerical solutions of the final equations only b is considered. Thus, it would be expected that the solution for 91 km would be most divergent from and yield larger values than the experimental data, therefore, providing a sort of theoretical check. Additionally, this procedure is adopted because uncertainties in b for 91 km is of the order of ε.

Our model incorporates a production and a loss term, together with initial conditions of thermal ionization as controlled by the initial expansion. The production term is affected by consumption and diffusion of the neutrals, but this neutral diffusion will only be effective at 91 km. The loss term for electrons includes the effect of ion–electron recombination, attachment, photodetachment and mutual neutralization; however, the diffusion is ignored. Our basic equation then is

$$\frac{dn_e}{dt} = \frac{kn_0}{1+\lambda}\, e^{-bt} - \alpha_e n_e^2 \tag{5}$$

where k = photoionization probability, $\sec^{-1}$
n_0 = neutral cesium density cm^{-3}
n_e = electron density cm^{-3}
b = chemical consumption probability $\sec^{-1}$
λ = ratio of negative ions to electrons
α_e = effective recombination coefficient

Further,

$$\alpha_e = \alpha_r + \lambda\alpha_i \tag{6}$$

where α_r = radiative recombination coefficient $\mathrm{cm}^3/\sec$
α_i = mutual neutralization (Cs^+ and O_2^-)

Although the non-linear equation (5) is susceptible to a complete formal solution, it is not susceptible to direct computation because of the large value of the arguments. However, numerical solutions have been obtained for equation (5) by the use of a variety of numerical approximations. The solutions, the approximations and their range of validity are included below.

The formal solution* to equation (6) can be written as

$$n_e(t) = \frac{bl}{2\alpha}\,\xi\,\frac{[I_1(l)K_1(l\xi) - I_1(l\xi)K_1(l)] + \mu[I_0(l)K_1(l\xi) + I_1(l\xi)K_0(l)]}{[I_1(l)K_0(l\xi) + I_0(l\xi)K_1(l)] + \mu[I_0(l)K_0(l\xi) - I_0(l\xi)K_0(l)]} \tag{7}$$

for all values of time $t \geqq 0$,

where $l = \frac{2}{b}\left(\frac{2kn_0}{1+\lambda}\right)^{\frac{1}{2}}$; $\mu = \frac{2\alpha n_e(0)}{lb}$ and $\xi = e^{-\frac{1}{2}bt}$.

It should be noted that the functions

$$y_1 = I_0(lZ) \text{ and}$$
$$y_2 = K_0(lZ)$$

satisfy Bessels' zero order modified equation,

$$\frac{d^2y_n}{dZ^2} + \frac{1}{Z}\frac{dy_n}{dZ} - l^2y_n = 0, \quad n = 1, 2, \tag{8}$$

and also that

$$I_0'(Z) = I_1(Z), \text{ where}$$
$$K_0'(Z) = -K_1(Z). \tag{9}$$

For application to the present data, the following approximations were utilized for the appropriate cases.

For times such that $l\xi = l\exp(-\frac{1}{2}bt) > 15$ the following equations apply

$$n_e(t) = \frac{bl}{2\alpha}\,\xi\coth[\coth^{-1}\mu + (1-\xi)l] \text{ for } \mu > 1 \tag{10}$$

$$n_e(t) = \frac{bl}{2\alpha}\,\xi\tanh[\tanh^{-1}\mu + (1-\xi)l] \text{ for } \mu < 1 \tag{11}$$

For times $1 < l\xi < 15$, $l > 15$, the following approximation holds

$$n_e(t) = \frac{bl}{2\alpha}\,\xi\,\frac{e^XK_1(X)}{e^XK_0(X)} = \frac{bl}{2\alpha}\,\xi\left(1 + \frac{0{\cdot}375}{X} - \frac{0{\cdot}1172}{X^2} + \frac{0{\cdot}10253}{X^3} - \right)\Big/ 1 - \frac{0{\cdot}125}{X} + \frac{0{\cdot}0703}{X^2} - \frac{0{\cdot}0732}{X^3} + \ldots. \tag{12}$$

* This mathematical solution together with the approximations are due to Dr. Herbert K. Brown, Geophysics Corporation of America.

where $X = l\xi$

Finally, if $l\xi < 1$, $l > 15$ the following expression is valid

$$n_e(t) = \frac{b/\alpha}{bt - 2\ln n} \tag{13}$$

It should be noted that the solutions represent the time-behavior at a single point. In this case, the comparison to the data was restricted in application to the center-point density. However, this restraint was not necessary since these solutions are applicable to the off-center points as well, including the case for the model employed here. This is occasioned by the fact that all the considered processes are purely time-dependent. Finally, this equation becomes increasingly less applicable with altitudes above approximately 80 km.

4.7. *Analysis of the* 40·7 Mc *radar data*

In section 3·7 it was reported that the observed 40·7 Mc back-scatter echoes were obtained for totals of 70 sec, 4·3 sec and 1·5 sec for the 91 km, 82 km and 69 km releases, respectively. The corresponding theoretical values of about 190 sec, 16 sec and 0·5 sec are obtained from Fig. 1. The three curves shown are plotted in accordance with equation (7) and show the variation of the center-point density with time. The parametric values employed are listed in Table 2.

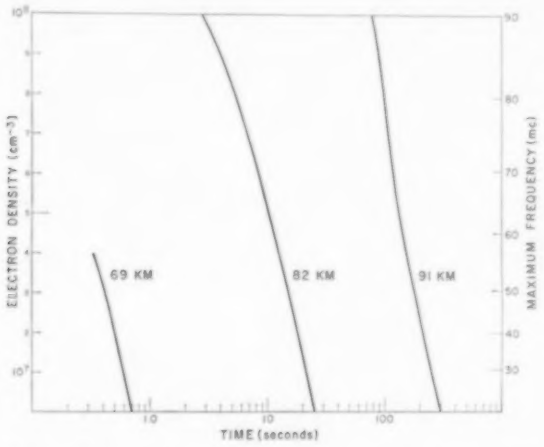


Fig. 1. Theoretical time variation of center-point electron density in accordance with equation 8a for 69, 82 and 91 km.

Table 2. Values of Basic Parameters Initially Used

H (km)	n_h (cm^{-3})	D_a ($cm^2\ sec^{-1}$)	D_i ($cm^2\ sec^{-1}$)	r_0 (cm)	n_0 (cm^{-3})	n_{e_0} (cm^{-3})	λ	α ($cm^3\ sec^{-1}$)	b (sec^{-1})
69	$2{\cdot}0\times10^{15}$	$5{\cdot}4\times10^{3}$	$6{\cdot}5\times10^{2}$	$3{\cdot}8\times10^{5}$	$3{\cdot}3\times10^{12}$	$6{\cdot}6\times10^{9}$	2	2×10^{-7}	8
82	$2{\cdot}5\times10^{14}$	$4{\cdot}0\times10^{4}$	$4{\cdot}0\times10^{3}$	$6{\cdot}9\times10^{5}$	$5{\cdot}5\times10^{11}$	$1{\cdot}1\times10^{9}$	0·2	2×10^{-8}	0·2
91	$4{\cdot}0\times10^{13}$	$1{\cdot}7\times10^{5}$	$1{\cdot}8\times10^{4}$	$1{\cdot}17\times10^{6}$	$1{\cdot}1\times10^{11}$	$2{\cdot}2\times10^{8}$	0·017	2×10^{-9}	0·017

n_h = number density at altitude.

n_{e_0} = initial center point electron density.

Other symbols have been defined in the text.

It should be noted that the general nature of the altitude variations and the absolute time values shown in Fig. 1 are in reasonable agreement with the observed 40·7 Mc radar data. The greatest divergence is noted for the 91 km case as expected from earlier discussions, since no allowance was made on the production term for the effect of diffusion. For example, if a correction is applied in this direction by utilizing b' with the previously derived values of ε, the estimated 40·7 Mc duration is reduced to only 100 sec. Additionally, the b-value for Cs may be larger than the value computed for Na. Finally, any gross error in the r_0 estimate would influence the pertinent time constant.

A similar case prevails for the 82-km release for which again, the computed value is larger than the observed value. Although it is felt that no significant allowance need be made for diffusion effects for altitudes as low as 82 km the b-value and r_0-value arguments given above for the 91 km case also apply here, but to a lesser degree.

Finally, unlike the other cases discussed, for the 69-km release the time behavior of the theoretical curve in Fig. 1 is relatively insensitive to the estimates adopted for b and r_0. For example, chemical consumption is so fast at this lower altitude that there is very little photoionization production and the governing process is electron attachment with subsequent mutual neutralization. Thus the radar echo time behavior is primarily dictated by the value employed for the product, $\lambda\alpha_i$. With λ-values taken from previous work[1] the α_i-values obtained by this analysis vary between 10^{-7} cm^3/sec to 10^{-8} cm^3/sec. In view of the estimated nature and values of the parameters utilized, the overall agreement between the observed and theoretical data is considered to be surprisingly good.

5. CONCLUSIONS

The analysis of the series of experiments discussed above has established a lower limit at about 70 km, below which long-lived, effective artificial electron clouds cannot be produced by the experimental techniques employed. This limit is due to the increasingly important roles of electron attachment and mutual neutralization with decreasing release altitudes. On the other hand, with higher altitude releases (up to 90 km) the electron cloud lifetime becomes increasingly sensitive to variations in b, the chemical consumption rate, and r_0, the half-width of the initial distribution. In addition, at 90 km and above, the effect of diffusion becomes important upon the photo-production rate. It should be noted that throughout the entire altitude region discussed here (70–90 km), ambipolar diffusion effects do not significantly affect the time variation of the center-point electron density.

The mathematical model employed seems reasonably adequate. Its apparent marked validity for application at the lower altitudes should appropriately be checked by more extensive data. The derived values of λ, α_r, α_i, b and k from this analysis seem to be of the correct order-of-magnitude. In particular for Cs, the value of b, a factor of two higher than that computed for Na, seems more appropriate. The fit also seems to support reasonably high values for α_i. However, while $\alpha_i = 10^{-7}$ cm^3 sec^{-1} appears to be the best value to adopt, a value of $\alpha_i = 10^{-8}$ cm^3 sec^{-1} also is compatible with this work.

Several experiments are suggested by this preceding work. At sufficiently low altitudes, errors in b and r_0 do not significantly affect the data, thus night time low altitude releases should give unambiguous values for electron attachment. The corresponding daytime low-altitude experiments should yield the value of the product $\lambda\alpha_i$. From the combined data it should be able to obtain more accurate values of α_i, the coefficient of mutual neutralization.

REFERENCES

1. F. F. Marmo, J. Pressman and L. M. Aschenbrand, *Planet. Space Sci.* **1**, 291 (1959).
2. F. F. Marmo, L. M. Aschenbrand and J. Pressman, *Planet Space Sci.* **1**, 306 (1959).
3. J. Pressman, F. F. Marmo and L. M. Aschenbrand, *Planet Space Sci.* **2**, 17 (1959).
4. F. F. Marmo, J. Pressman, E. R. Manring and L. M. Aschenbrand, *Planet Space Sci.* **2**, 174 (1960).

5. Harvard University Solar Laboratory, private communication.
6. R. T. BURNIGHT, *Ultraviolet Radiation and X-Rays of Solar Origin* in *Physics and Medicine of the Upper Atmosphere* (edited by C. S. WHITE and O. O. BENSON) p. 227. University of New Mexico Press, New Mexico (1952).
7. D. R. BATES, *The Physics of the Upper Atmosphere* in *The Earth as a Planet* (edited by G. P. KUIPER) p. 618. University of Chicago Press, Chicago (1954).
8. S. K. MITRA, *The Upper Atmosphere*, p. 313. The Asiatic Society, Calcutta (1953).
9. A. DALGARNO, private communication.

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 238–245. Printed in Great Britain.

A STUDY OF "CHORUS" OBSERVED AT AUSTRALIAN STATIONS

J. CROUCHLEY and N. M. BRICE*

Physics Department, University of Queensland, Brisbane

(*Received* 14 *October* 1959)

Abstract—The very-low-frequency emission known as "chorus" has been studied for four Australian recording stations. The phenomenon exhibits a maximum of occurrence at about geomagnetic latitude 60° and occurs only about one-tenth as frequently at geomagnetic latitude 45°. It was never definitely observed at geomagnetic latitude 35°. Marked diurnal variations are present with morning peaks for all stations and also an evening peak at Adelaide and Hobart. Analysis of the diurnal variations suggests that two different mechanisms are involved. The average strength of chorus increases with increasing value of *K*-index at lower latitude stations and shows a maximum at moderate values of the index for geomagnetic latitude 61°. This is interpreted as meaning that the region of maximum average-chorus-strength moves towards lower latitudes at times of magnetic disturbance.

1. INTRODUCTION

Storey[1] reported the reception, on an aperiodic very low frequency radio receiver, of signals which on a loudspeaker sounded like "a rookery heard from a distance". More precisely they may be described as a mixture of rising and falling tones and/or quasi-stationary tones of a frequency of a few kilocycles per second. The duration of an individual tone varies between a fraction of a second and a few seconds, and the electric field strength may be of the order of a few microvolts to some millivolts per metre. The phenomenon, which had been observed previously as interference on audiofrequency communication circuits, was originally named "dawn chorus" owing to its having a maximum occurrence about local sunrise but since subsequent investigations have shown that the time of this maximum depends upon latitude the more recent term "chorus" will be used here. The general term "Very-low-frequency emissions" includes also the longer-lasting and more uniform emission known as "hiss"; the latter has been disregarded in the work to be described.

Gallett[2] and other workers have suggested that V.L.F. emissions are caused by the interaction of fast-moving particles—originally derived from the sun—with the upper atmosphere and the magnetic field of the earth, and it has been proposed by Allcock[3] that the radiation travels along magnetic lines of force in the "whistler mode". Travelling-wave amplification[2, 4, 5] and Cerenkov radiation[6, 7] have been suggested as generating mechanisms of V.L.F. emissions.

An association of the phenomenon with magnetic disturbance was suggested by Storey[1] and evidence for this was presented by Allcock[3].

2. EXPERIMENTAL DETAILS

As an International Geophysical Year project, the University of Queensland installed equipment for recording whistlers at Brisbane (geographic lat. 27°32′S., long. 152°55′ E.; geomagnetic lat. 35°8′S, long. 226°9′), Adelaide (34°57′, 138°23′; 44°9′, 212°5′), Hobart (42°52′, 147°20′; 51°7′, 224°6′) and Macquarie Island (54°30′, 158°57′; 61°1′, 243°1′). The recording equipment had a frequency range of from 400 c/s to 15 kc/s and was operated automatically for 2-min periods each hour commencing at 35 min past the hour (Universal Time). The Brisbane, Adelaide and Hobart recorders were of the same design and used a

* Now at Department of Electrical Engineering, Stanford University California.

12-turn square loop aerial of 5 m side. It is estimated that a field strength of 10 to 15 μ V/m would give detectable signals on these recorders in the absence of external noise. External noise was however almost always the limiting factor. At Macquarie Island two different recording equipments were used. The first equipment (hereafter referred to as the *B*-recorder) used a 10 m vertical aerial and an aerial preamplifier of somewhat lower gain than that at the other stations. It is estimated that the sensitivity of this recorder was one-half to one third of that of the recorder at Hobart. The *B*-recorder was replaced in February 1958 by one (hereafter referred to as the *A*-recorder) kindly supplied by Dr. R. A. Helliwell of Stanford University. Chorus strength was graded on a subjective aural basis from 0–5. It should be pointed out that the detection and assessment of the phenomenon, particularly when weak, depend very markedly upon the general noise level (including atmospherics and man-made). Thus no great precision attaches to the threshold of detection or the aural grading at low strengths.

3. VARIATIONS IN CHORUS ACTIVITY

(a) *Diurnal variation*

Fig. 1 is a plot of the average percentage occurrence of chorus within each sampling period of the day against time of day (local mean time). Adelaide and Hobart data were averaged from mid-1957 to mid-1958. In the case of Macquarie Island the averages for the *B*-recorder (1957) indicated a similar diurnal variation to those of the *A*-recorder (1958) but the latter were systematically higher by a factor of about three. To simplify presentation the *A*-averages were therefore divided by three and the mean between these and the *B*-averages shown on the diagram. The phenomenon was never definitely observed in Brisbane although very occasionally faint signals of doubtful authenticity were heard.

Macquarie Island shows a marked broad maximum of occurrence a few hours before local midday, Hobart a similar but narrower maximum somewhat earlier in the day and a smaller peak in the early evening. Adelaide shows two broad peaks, morning and evening and a lesser peak around midday, but, as the total number of occurrences for each sampling period was less than twenty, no great significance can be attached to the minor features of the curve, especially the midday peak which has no counterpart on the plots for Hobart or Macquarie Island.

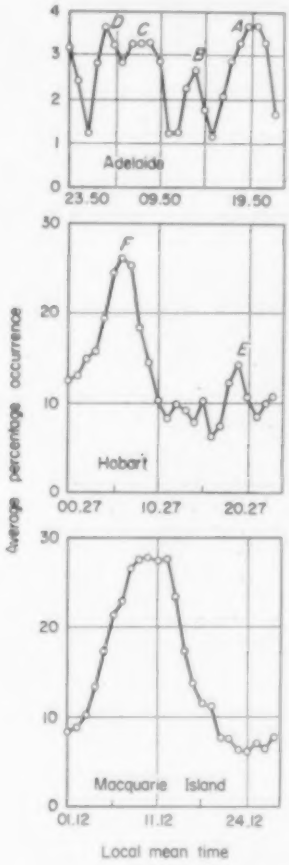


Fig. 1. Diurnal variation.

The results suggest that at Adelaide the probability of occurrence has equal morning and evening peaks, at Macquarie Island only a morning peak, while at Hobart, which is intermediate in geomagnetic latitude, there are two

peaks, the morning one predominating. In an attempt to assess whether the difference in behaviour indicates different mechanisms for morning and evening chorus, groups of sampling periods representing the different peaks in Adelaide and Hobart were taken. These groups are set out in Table 1 (letters refer to Fig. 1). Calculations were made of the correlation coefficients between the monthly totals of the groups and are shown in the second column of the Table.

The comparison between Hobart and Macquarie Island was complicated by the change of recorder at Macquarie Island and accordingly the two periods were treated separately and then combined using a scaling factor of three as above. As shown by Table 2 most of the correlation coefficients are small and below the 10 per cent significance level. There is, however, a suggestion of a weak negative association and more so for the Hobart morning peak than the total occurrence.

Table 1. Adelaide

Hours in group (all inclusive)	Correlation coefficient (r) between groups	No. of pairs of values	Value of r for significance at %-level
A=19-22	A-B 0·64	14	0·66 1%
B=13-14	A-C 0·73	14	0·66 1%
C= 7- 9	A-D 0·72	14	0·66 1%
D= 3- 5			
x=all hours			
	Hobart		
E=18-19	E-F 0·47	16	0·50 5%
F= 5- 7			
y=all hours			
	Adelaide vs. Hobart		
	x-y 0·51	14	0·53 5%
	E-x 0·52	14	0·53 5%
	A-E 0·59	14	0·53 5%
	D-F 0·10	14	0·53 5%
	C-F 0·09	14	0·53 5%
	A-F 0·00	14	0·53 5%
	x-F 0·06	14	0·53 5%

Table 2. Hobart and Macquarie Island

	Correlation coefficient	Number of pairs of values	Value of r significant at 10%-level
Hobart 1957 (all hours) B recorder (all hours)	−0·37	9	0·58
Hobart 1958 (all hours) A recorder (all hours)	0·21	8	0·62
Hobart (all hours) B recorder and 1/3 A recorder	−0·02	17	0·41
Hobart F-B recorder (all hours)	−0·51	9	0·58
Hobart F-A recorder (all hours)	−0·20	8	0·62

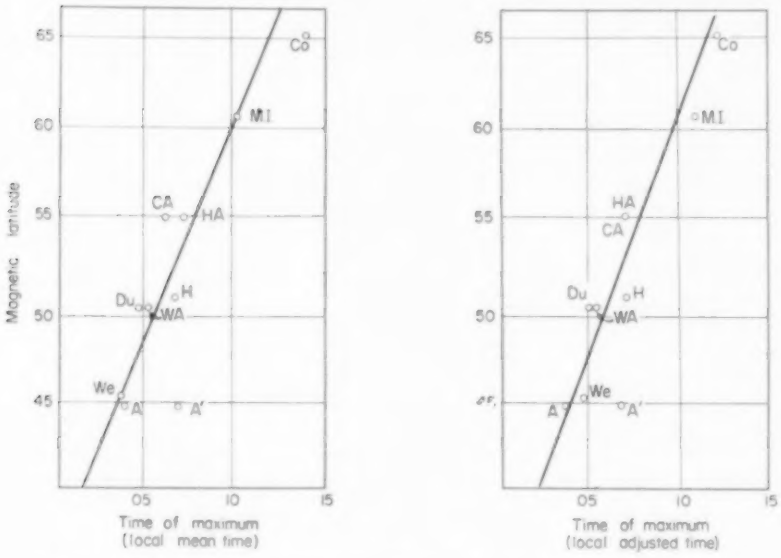


Fig. 2. Variation of time of maximum of "chorus" with geomagnetic latitude.
(A—Adelaide (peak D), A'—Adelaide (middle of C and D peaks), CA—Cambridge (U.K.), Co—College (Alaska), Du—Dunedin (N.Z.), H—Hobart, HA—Hanover (U.S.A.), M.I.—Macquarie Island, WA—Washington (U.S.A.), We—Wellington (N.Z.))

(b) *Variation of time of maximum occurrence with geomagnetic latitude*

Pope[8] has investigated the variation with geomagnetic latitude of the local mean hour for maximum occurrence of chorus. Fig. 2(a) indicates that the Macquarie Island and Hobart (morning) peaks fit well into the general world picture. If *D* (Fig. 1) is regarded as the morning peak for Adelaide, there is again good agreement, but if *D* and *C* are regarded as parts of a single broad maximum the centre is well off the Pope curve. Evening peaks have not been reported at stations in other parts of the world.

If the generation of chorus is controlled by the injection of particles at the top of a line of force leading to the observing station, then the transit of the sun above this region may be a more significant time reference than its transit over the geographic meridian. Times of maxima were therefore adjusted to allow for the differences between geographic and geomagnetic meridians at the geographic equator. Fig. 2(b) shows the data of Fig. 2(a) replotted with appropriate adjustment. The fit of the points to a straight line is slightly better and the discrepancy in the case of Adelaide is reduced.

(c) *Day to day variation*

In order to estimate the duration of bursts of chorus and also to search for any recurrence tendency in the data auto-correlation analyses were performed on 24-hr totals, weighted in this case by the strength indices. The start of the 24-hr period was adjusted to fall in or near a minimum of the diurnal distribution for the station in question, e.g. midnight for Hobart.

For Hobart the auto-correlation coefficients were significantly above zero for one-day and two-day lags and thereafter (up to 65 days lag) showed little systematic behaviour (though there was a general tendency for small negative values for lags between 30 and 55 days). For Macquarie Island *A*-recorder data, where chorus was a common occurrence, significant values of correlation coefficient were obtained

with delays up to about one month. It was felt, in this case, that an analysis of departures from a running-average centred on the day in question (13 days before to 13 days after) would be more indicative of changes in chorus activity. This analysis showed small but significant values of correlation coefficient for 1-day and 2-day lags. Auto-correlation analysis of the Adelaide data, where on most days chorus was not observed, was not considered to be practicable.

The general conclusion is that chorus activity has no recurrence tendency at Hobart and Macquarie Island (in particular no 27-day recurrence tendency) and that a period of above average strength tends to last 1–2 days.

(d) *Variation with geomagnetic latitude*

As mentioned earlier, the aural detection of chorus is to a marked degree dependent on local noise conditions and on the equipment used. The comparison of Hobart and Macquarie Island in particular is complicated by the difference in listening conditions (quieter at Macquarie Island) and recorder characteristics. There seems however no doubt that geomagnetic latitude 44° shows a markedly lower incidence (3 per cent yearly average) than latitude 51° (14 per cent) and latitude 61° (15 per cent *B*-recorder, 46 per cent *A*-recorder). Taking into account the difference in conditions between Hobart and Macquarie Island, it seems likely that the average occurrence at latitude 61° is about twice that at latitude 51°.

Pope's[8] results for a different year suggest that the incidence chorus at geomagnetic latitude 65° is appreciably less than for latitudes 60°. Martin[9] did not observe chorus at geomagnetic latitude 79° during a period of observation of a fortnight although whistlers were observed. This lack of occurrence would have been very unlikely if the actual average occurrence were comparable to that at Macquarie Island.

The region of maximum activity thus seems to be around 60° geomagnetic latitude or possibly a few degrees less and the occurrence rate seems to fall off more rapidly on the polar side of this latitude than on the equatorial side.

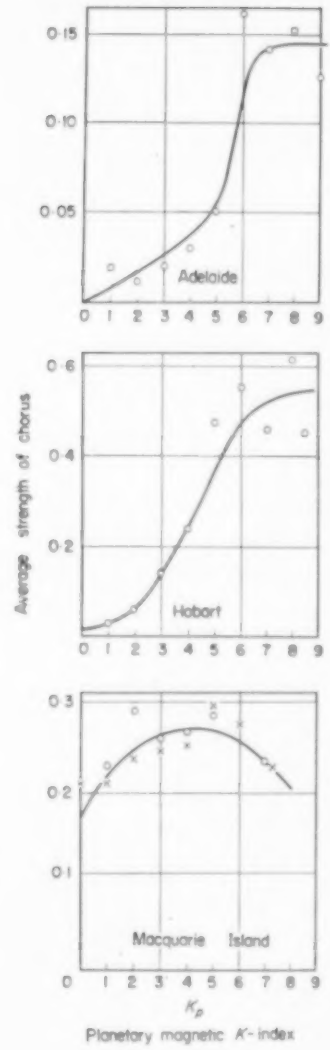


Fig. 3. Relationship between chorus strength and planetary magnetic index.

4. ASSOCIATION OF CHORUS WITH MAGNETIC DISTURBANCE

Fig. 3 shows a plot of average chorus strength (no occurrence being counted as zero strength) at each station for each value of planetary magnetic K-index (K_p) at the relevant time the data being averaged over one year. In

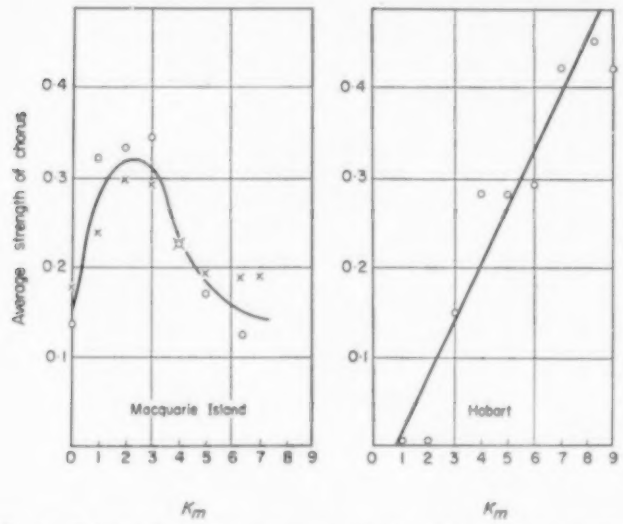


Fig. 4. Relationship between chorus strength and local magnetic index.

the case of Macquarie Island *B*-recorder data has been plotted with a dot and *A*-recorder data with a cross and scale reduction factor of three. As high values of the *K*-index occur relatively infrequently the Macquarie Island data has been grouped for K_p greater than five. The Adelaide and Hobart curves are not far from parabolic between $K_p=0$ and $K_p=6$ but give indications of "saturation" at higher values while Macquarie Island which shows a much greater average strength at magnetically-quiet periods shows also indications of a decrease in strength above $K_p=5$.

The contrast between Hobart and Macquarie Island is even better brought out in Fig. 4 which shows a similar plot of chorus at these stations against the *K*-index (K_m) for Macquarie Island. (Magnetic observations were not made at Hobart.) The maximum chorus strength occurs about $K_m=3$ for Macquarie Island but the strength at Hobart increases steadily with rising K_m. The very low chorus strength at Hobart for K_m less than 3 is also striking and markedly different to Macquarie Island which station has an average-strength about one-half of the maximum average-strength for $K_m=0$.

The Macquarie Island 27-day running-average strength figures were compared with similar data for K_p as shown in Fig. 5 where the values obtained from K_p data have been plotted

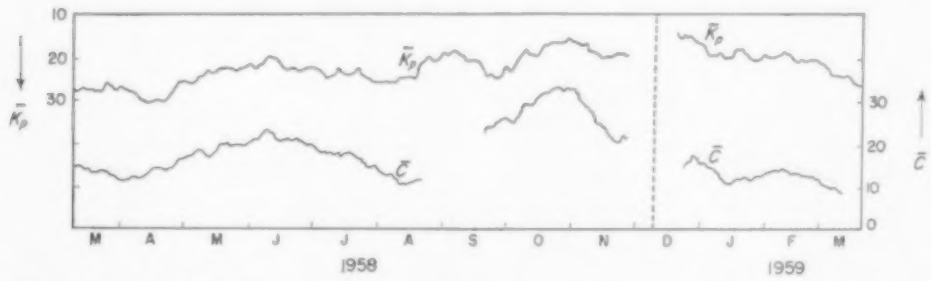


Fig. 5. Comparison of planetary magnetic index 27-day running averages ($\bar{K}p$) and Macquarie Island chorus strength 27-day running averages ($\bar{C}$) 28 days later.

increasing downwards and also displaced to a date 28 days later. (The first break in the strength curve occurs because of equipment failure and the second because of a change of observer.) There is sufficient similarity between the two curves to suggest that consistently high values of K_p, i.e. magnetically disturbed periods lasting some days, tend to reduce the average strength of chorus at this station for a corresponding period some weeks later.

5. RELATION TO SOLAR PHENOMENA

As it has been suggested that the primary generating mechanism of chorus is associated with the arrival of solar particles a comparison was made of solar and chorus data. No association could be found with sunspot number, solar flares or Type IV radio flares. However some indication of correlation was found between occurrence in Adelaide and rate of change of sunspot number. For this purpose the period of observations was divided into intervals of 27 days and the change in average sunspot number from one interval to the next determined. This was compared with the percentage occurrence of chorus in a 27-day interval centred on the day dividing the appropriate two intervals used for sunspots. Over fifteen consecutive periods the Adelaide data showed a significant (1 per cent level) positive correlation between chorus occurrence and increase in sunspot number but the inclusion of the sixteenth pair of values from data at the end of 1958 reduced the correlation coefficient to one significant at about a 7 per cent level. Hobart and Macquarie Island data did not suggest a similar relationship.

6. DISCUSSION

The differences in the diurnal variation exhibited by the three stations together with the indications (shown in Table 1) of positive statistically significant correlation between morning and evening peaks in Adelaide, less association between morning and evening peaks in Hobart, weak association between Adelaide and Hobart evening peaks, very little association between Hobart morning peak and Adelaide and a weak negative association between Hobart morning peak and Macquarie Island suggests that two somewhat distinct mechanisms may be responsible for the occurrence of chorus. One of these (i) may be considered to be operative roughly equally morning and evening at Adelaide and Hobart stations and the other (ii) only in the morning and at Hobart and Macquarie Island. In this connection it may be noted that in order to explain the observed minimum around midday in the occurrence of whistlers at moderate latitudes[10, 11] it has been postulated that ionospheric absorption during daylight hours prevents the detection of weaker whistlers. If we assume a mechanism for generation of chorus equally effective for the whole daylight period (possibly displaced by up to 2 hr because of skewness of magnetic lines of force) we should then expect reception for two roughly equal periods, one just after sunrise (at the top of the line of force) and the other just after sunset (at the same place). This accords with mechanism (i) but not mechanism (ii) for which the generating mechanism would seem to have a peak activity before midday.

The relationships between chorus strength and magnetic disturbance indices indicate that, at least for the mechanism (ii), the latitude peak of occurrence—which, for low K-values, is near the latitude of Macquarie Island—is displaced at times of great magnetic disturbance nearer to the latitude of Hobart. Since at these times the chorus strength and hence probability of detection is reduced at Macquarie Island and increased at Hobart, this shift is consistent with the suggestion of negative correlations between Macquarie Island and Hobart (morning peak) occurrences. If it is assumed that at any one time chorus is confined to a geomagnetic latitude zone about 10° wide which is centred somewhere near geomagnetic latitude 63° for very magnetically quiet times ($K_m=0$) and shifts about 1° towards the equator for each increase of 1 in K_m then, from the fact that K_m was over 3 for 43 per cent of the time and 0 for only 6 per cent, we can roughly account for the observed latitude dependence of chorus occurrence.

Assuming a dipole field, the line of force passing through the maximum of the upper Van Allen radiation belt[12], as at present plotted, would reach the surface at a geomagnetic latitude between Hobart and Macquarie Island. If the spiralling charged particles in this belt[13] are responsible for the generation of chorus one would then expect a maximum occurrence in this region. On this hypothesis the movement towards the equator, of the region of maximum chorus activity during times of magnetic disturbance would suggest a downward movement of the generating region. Similar movements have been postulated by Obayashi[14] as an explanation of geomagnetic pulsation phenomena. Obayashi's analysis shows the upper boundary of the earth's atmosphere descending to as low as two earth radii from the centre, i.e. approximately the height of the top of a geomagnetic line of force through Adelaide, for brief periods. A comparison of geomagnetic pulsation and chorus records (preferably taken much more frequently than in the records for this study) might well facilitate the explanation of the latter phenomena.

Acknowledgments—Acknowledgments are due to the Australian National Committee of the International Geophysical Year and the University of Queensland for the provision of finance for equipment, to the Antarctic Division Department of External Affairs, the Ionospheric Prediction Service, Hobart, and Weapons Research Establishment, Adelaide, for the housing and operation of equipment and to the several operators, particularly Mr. R. Hosking and Mr. B. McInnes who listened to many tapes. The *B*-recorder records were made while one of us (N.M.B.) was with the Australian National Antarctic Research Expedition. The information concerning solar radio emissions was kindly supplied by the Radiophysics Laboratory of C.S.I.R.O. It is a pleasure to acknowledge the unfailing interest and assistance of Professor H. C. Webster.

REFERENCES

1. L. R. O. Storey, *Phil. Trans. Roy. Soc.* **246,** 113 (1953).
2. R. M. Gallet, *Proc. Inst. Radio Engrs., N.Y.* **47,** 211 (1959).
3. G. McK. Allcock, *Aust. J. Phys.* **10,** 286 (1957).
4. R. A. Helliwell, *Low Frequency Propagation Studies,* part 1, AFCRC Techn. Rep. 56-189 (1956).
5. R. M. Gallet and R. A. Helliwell, *J. Res. Nat. Bur. Stand.* **63D,** 21 (1959).
6. G. R. Ellis, *J. Atmos. Terr. Phys.* **10,** 302 (1957).
7. N. M. Brice, to be published.
8. J. H. Pope, *Nature, Lond.* **180,** 433 (1957).
9. L. H. Martin, *Nature, Lond.* **181,** 1796 (1958).
10. R. A. Helliwell and M. G. Morgan, *Proc. Inst. Radio Engrs., N.Y.* **47,** 200 (1959).
11. J. Crouchley, to be published.
12. J. A. van Allen and L. A. Frank, *Nature, Lond.* **183,** 430 (1959).
13. J. A. van Allen, C. E. McIlwain and G. H. Ludwig, *J. Geophys. Res.* **64,** 271 (1959).
14. T. Obayashi, *Rep. Ionos. Res. Japan* 12 (1958).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 246–248. Printed in Great Britain.

THE TIME DELAY OF THE RED [OI] LINES IN THE AURORA

A. OMHOLT

The Institute of Theoretical Astrophysics and the Institute of Physics, University of Oslo

(*Received* 15 *December* 1959)

Abstract—A detailed numerical analysis of Stoffregen and Derblom's records of the red line λ 6300 and the N_2^+ band λ 4278 have been undertaken. The red line is delayed relatively to the violet band, the average delay time being 150 to 200 sec. The average intensity ratio $I(6300)/I(4278)$ is about 4·6 for these records. These results may have important consequences for the theory of the excitation and deactivation of the $O(^1D)$ atoms, which is discussed.

Stoffregen and Derblom[1] have obtained some very interesting time records of the intensities of the red [OI] line λ6300 and the N_2^+ band λ4278 emitted from aurorae. They found a displacement between the two intensity curves, the red line being delayed relatively to the N_2^+ band. They attributed this delay to the long lifetime of the metastable $O(^1D)$ atoms, and found a mean value of about 140 sec for this delay, measured as the average displacement between the two intensity curves along the time-axis.

The twelve records analyzed by Stoffregen and Derblom, together with five more, also obtained by them, have been analyzed by a linear, numeric method which was used in a similar study of the green [OI] line λ5577, described in an earlier paper.[2] This method is based upon the assumption that the ratio between the two rates of excitation as well as the average lifetime τ of the $O(^1D)$ atoms are constants for any one record, this lifetime being alone responsible for the observed time delay. The purpose of this note is to compare the two methods and to discuss the physical interpretation of the results.

τ as well as the average intensity ratio $I(6300)/I(4278)$ were computed, and the results are given in Table 1, together with the values obtained by Stoffregen and Derblom. They distinguish between the delay times found from the rising (t_r) and the falling (t_f) part of the curves respectively, and find that t_f is usually larger than t_r. This difference is much due to the scaling of the curves. If these are scaled to give equal integrated amplitudes (i.e. equal average amplitudes) rather than equal peak amplitudes, the difference becomes insignificant. Record numbers 13–17 were not included in

Table 1. Computed Lifetimes (τ) and Average Intensity Ratios ($I(6300)/I(4278) = k$). t_r and t_f are Stoffregen and Derblom's Values for the Delay from the Rising and Falling Parts of the Curves

Rec. no.	k	τ (sec)	t_r (sec)	t_f (sec)
1	2·0	110	99	168
2	1·8	175	102	195
3	2·1	205	105	222
4	0·3	55	115	109
5	4·2	525	120	180
6	7·9	130	127	330
7	3·3	120	53	135
8	13·7	120	75	180
9	2·8	140	135	120
10	8·5	135	135	135
11	2·3	160	135	160
12	0·7	20	135	0*
13	5·9	725		
14	0·7	300		
15	7·8	145		
16	3·6	−10		
17	10·0	215		
Mean 1—12	4·1	160	110	160
Mean 1—17	4·6	190	—	—

* This value has been added by the author, with a corresponding change in the mean.

Stoffregen and Derblom's analysis, because they are rather complicated and not easy to analyse by their method. The mean value of τ from the first twelve records is 160 sec and from all seventeen records 190 sec. Both these values are somewhat greater than the mean value obtained by Stoffregen and Derblom, but the difference cannot be regarded as significant. Thus the apparent lifetime is somewhat larger than the theoretical radiative lifetime of the $O(^1D)$ atoms, which is about 110 sec.[3]

The mean of the average values of the intensity ratio $I(6300)/I(4278)$ (k) is about 4·6, and the individual values vary between 0·3 and 13·7. Since $I(5577)/I(4278)$ is about three—cf.[2] compared with[4]—the mean value of $I(6300)/I(5577)$ should be about 1·5 for the observed aurorae. This value is much less than the computed ratio between the excitation rates of the two oxygen lines, which is greater than 10 [5, 2]. There is no significant correlation between τ and k.

Thus the average of the measured time delay between the λ4278 band and the λ6300 line is rather larger than the radiative lifetime of the $O(^1D)$ atoms whereas the average of the intensity ratio $I(6300)/I(4278)$ is lower than the expected value by a factor 5 or more. There may be at least two possible explanations for this. If we assume that the low value of the intensity ratio is due to collisional deactivation of the $O(^1D)$ atoms, then the average lifetime for these atoms in an aurora should be of the order 10 sec rather than 100 and the observed time delay between the two emissions largely due to the excitation processes. On the other hand, if we assume that the measured average delay time is close to the radiative lifetime of the $O(^1D)$ atoms, collisional deactivation being negligible, then the excitation processes must differ from those thought to be operative.

If the first explanation is correct there must be a cause for the time delay between the two excitation processes. If dissociative recombination of O_2^+ contributes appreciably to the excitation of the red lines, this may be such a cause (cf. discussions in[2]). Also it may be thought that since the collisional quenching of the red line is most effective at low altitudes, and the observed λ6300 light therefore originates at relatively high levels in the atmosphere, the excitation at high altitudes may be delayed relatively to that at low altitudes. This could again be due to a later arrival of the slower electrons which presumably excite the higher part of the aurora. The lowest energy a primary electron may have if it shall be able to excite the λ6300 line is 2 eV, corresponding to an electron velocity of about 10^8 cm/sec. If the electrons exciting the lower and higher parts of these bursts of bright aurora have a common origin in space and time, this origin should then be at a distance 10 to 100 earth radii from the auroral zone.

If the second explanation is correct, collisional deactivation of the $O(^1D)$ atoms is unimportant. The low rate of excitation of the red line may be explained by assuming that other processes compete with the oxygen atoms for the energy of the electrons. Thus most of the electrons in the 2–4 eV range, which are the most efficient ones in exciting the λ6300 lines, may spend their energy on other processes, as proposed in an earlier paper.[2]

Now, we must also bear in mind that our measurements do not necessarily yield the average lifetime of all the $O(^1D)$ atoms. What we measure is rather a weighted average, with preference on light from the regions in the atmosphere where the deactivation is least. To illustrate this, imagine the combined emission from two boxes, one where the $O(^1D)$ atoms are undisturbed and one where they are all deactivated by collisions. The measured lifetime would be that of the atoms in the first box, which are the only ones radiating the λ6300 line, whereas the measured intensity ratio would be influenced by the collisional deactivation in the other box. If the radiative lifetime for the $O(^1D)$ atoms were about 250 sec or so, a similar explanation of the auroral observations might be possible, but the observed height distribution curves[6] leave little doubt that the measured time delay would be significantly shorter than the radiative lifetime.

In any case the relationship between the in-

tensity curve for the λ6300 line and the λ4278 band is expected to be rather complex, since the ratio between the two excitation rates may vary considerably with height and time. The scatter in the apparent lifetimes of the O(^1D) atoms is therefore not surprising. A closer examination of the measured curves indicates also that our original assumption (constant excitation ratio and lifetime) is not correct. Therefore it may be somewhat doubtful which method of analysis is correct, but it is unlikely that the delay time is erroneous by a factor of more than 1·5, whatever its interpretation may be. This is also supported by the reasonable agreement between our results and those obtained by Stoffregen and Derblom.

Acknowledgements—The author is greatly indebted to Mr. W. Stoffregen and Mr. H. Derblom for generously putting their records and results at his disposal in advance of publication. This work was supported in part by the Geophysics Research Directorate of the U.S. Air Force, Cambridge Research Center, Air Research and Development Command, under contract AF61(514)-1123 through the European Office, Brussels.

REFERENCES

1. W. Stoffregen and H. Derblom, *Nature, Lond.* **185**, 28 (1960).
2. A. Omholt, *Geofys. Publ.* **21**, no. 1 (1959).
3. R. J. Garstang, *Mon. Not. R. Astr. Soc.* **111**, 115 (1951).
4. A. Omholt, *J. Atmos. Terr. Phys.* **6**, 61 (1955).
5. M. J. Seaton, *Astrophys. J.* **127**, 67 (1958).
6. L. Harang, *Geofys. Publ.* **20**, no. 5 (1958).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 249–255. Printed in Great Britain.

STRESS HISTORY OF THE MOON

G. J. F. MacDONALD

(*Received* 6 *January* 1960)

Abstract—The character of the lunar surface indicates that surface faulting has not been an important mechanism for the build-up of the lunar surface. If the radioactive content of the Moon is of the same order as that of chondritic meteorites, then the absence of major surface faults can be explained in a number of ways. A near-surface concentration of radioactivity will provide an equality of heat production and surface heat flow necessary for the maintenance of a constant lunar radius. Alternatively, the radioactivity could be deeply buried, with the radius still remaining constant over the past 2,000,000,000 years. Heat transported by mechanisms other than radiation and thermal conduction will also tend to keep the radius of the Moon at a constant value.

Even though the radius of the Moon remains constant, there is a major build-up of strain energy throughout the Moon. The rate is such that, on the average, something on the order of 10^{24}–10^{25} ergs of distortional energy should be released per year throughout the Moon, provided the radioactivity is uniformly distributed. A near-surface concentration of the radioactivity might decrease this rate of energy release but certainly by no more than an order of magnitude. Under all circumstances it would appear that a Moon of chondritic composition would have strong seismic activity.

The recognition of large-scale strike-slip faults on the lunar surface should be far easier than in the case of terrestrial strike-slip faults. The lunar surface is free of masking sediments and oceans. Furthermore, the complicated character of the lunar surface provides an abundance of landmarks for the detection of relative horizontal movement.

Examination of the Moon through a large telescope reveals the surface to be covered with craters. These round, circular depressions range in diameter from about 1 km, nearly the limit of observation, to a diameter of some 240 km. The veritable lacework of craters provides a background for the detection of relative horizontal movement along the lunar surface. Any lateral displacement of the circular depressions could easily be detected, provided the horizontal movement was on the order of half a kilometer or greater. A detailed examination of the lunar surface reveals no such lateral displacement in the craters.

There are a number of linear features on the Moon's surface which further indicate the absence of any large-scale horizontal displacement. The rays extending from Tycho are straight and there is no indication of any horizontal movement in the area which they cover. Nor do other linear features, such as the ridges, show any evidence for horizontal displacement. In summary, there are no reports of horizontally displaced features in the works of students of the lunar surface.[1, 2]

The absence of horizontally displaced features on the surface of the Moon suggests that there are no large-scale strike-slip faults or wrench faults on the Moon. There are a number of examples of faults on the earth's surface in which a relative horizontal movement of tens and perhaps hundreds of kilometers have taken place. The most famous of these is perhaps the San Andreas Fault in California. Recent magnetic studies in the ocean off the coast of California suggests a system of strike-slip faults with displacements of from 10 to 100 km. There are numerous, large strike-slip faults traversing the folds of the Jura Mountains. The maximum horizontal displacement along these faults is of the order of tens of kilometers. Kennedy[3] has suggested that there has been a displacement along the great Glen Fault of Scotland of more than 100 km. If faults of

this scale were present on the Moon, they should be easily visible, provided the faulting postdated the formation of the craters and other surface features of the moon.

Associated with collision craters are steep walls that can be interpreted as faults with vertical displacement. These faults are not expressive of the internal character of the Moon but are the results of the impact of meteorites on its surface. It has been suggested that certain of the great lunar valleys, such as the Rheita Valley, are valleys bounded by normal faults. The Rheita Valley is about 350 km long, 20–40 km wide and averages about 1·6 km deep. The Borda Valley is of similar dimensions. These valleys resemble the terrestrial African Rift Zone which extends some 600 km from north to south. The elongated, lunar depressions certainly are suggestive of a subsidence between normal faults, but there is no evidence of horizontal displacement.

Further indications of lunar faulting can be found in various apparent cracks on the lunar surface. Perhaps the most suggestive of these lies west of Copernicus where a number of small craters lie along a crack.

A study of the lunar surface therefore suggests that there may be normal faults with substantial vertical displacement and that faulting may have given rise to the great lunar valleys. There is no evidence whatsoever for large-scale strike-slip faults. The surface of the Moon is certainly not a mosaic of interconnected blocks separated by fault planes. If the lunar surface has ever undergone any extensive faulting, the faulting must have preceded the formation of the present surface features of the Moon.

The absence of evidence for extensive faulting on the lunar surface, and, in particular, the absence of evidence for large-scale strike-slip faults places limitations on the possible stress and thermal history of the moon. The lack of features associated with faulting implies that the volume and thus the surface dimensions of the Moon has remained more or less constant since the time of formation of the surface features of the Moon. If the circumference of the Moon has been increasing, then roughly speaking we should expect the presence of normal faulting and the formation of grabens or rifts. If the circumference of the lunar surface has decreased, then strike-slip faults and thrust faults would be expected from the resulting compression. A decrease in radius of 1 km of the Moon would result in a decrease in lunar surface area of $4{\cdot}5 \times 10^4$ km^2 or about one-tenth of one per cent of the present surface area.

The change of the radius of the Moon can result from various thermal processes. If the Moon were initially at a high temperature and the rate of loss of heat was greater than the rate at which heat was being produced in the interior, then the lunar radius should have decreased with time. If, on the other hand, the rate at which heat is being produced is greater than the rate at which heat is being lost by the Moon, then its radius should be increasing with time.

REVIEW OF CONTRACTION THEORY

We consider a spherical mass having a spherically symmetric density distribution. An equation for the conservation of mass is

$$\int_v \frac{\partial \rho}{\partial t}\, dv = -\int_v \frac{\partial}{\partial x_i}(\rho v_i)\, dv = -\int_S \rho v_r dS \qquad (1)$$

where ρ is the density, dv is the element of volume, and v_r is the radial velocity. Taking R as the outer radius of the spherical mass, the conservation of mass relation for a sphere is then

$$\frac{dR}{dt} = -\frac{1}{R^2 \rho_s}\int_0^R \frac{\partial \rho}{\partial t}\, r^2 dr \qquad (2)$$

where ρ_s is the density of the material at the outer surface.

We consider the effect of temperature on the density. In doing so we assume that the density changes only through thermal expansion. The possibility of phase changes and chemical reactions can be taken into account by use of an "effective thermal expansion" greater than the ordinary thermal expansion. The rate of change of density at any point

within the sphere is then related to the change of temperature at any point by

$$\frac{\partial \rho}{\partial t} = \left(\frac{\partial \rho}{\partial T}\right)\frac{\partial T}{\partial t} = -\rho\alpha\frac{\partial T}{\partial t} \qquad (3)$$

where α is the coefficient of thermal expansion.

The local change of temperature at a given point is determined by the divergence of the heat flux and the rate $A(r, t)$ at which heat is being produced per unit volume:

$$C_p\frac{\partial T}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial_r}\left(r^2K\frac{\partial T}{\partial t}\right) + A(r, t) \qquad (4)$$

where heat transported by mass motion is neglected. K in equation (4) is the effective conductivity of the material and may include contributions from radiation. Combining equations (2), (3) and (4) we obtain that the change of the outer radius with time is given by

$$\frac{dR}{dt} = \frac{1}{R^2\rho_s}\int_0^R \frac{r^2\alpha}{C_p}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2K\frac{\partial T}{\partial t}\right) + A(r, t)\right]dr \qquad (5)$$

Assuming a constant thermal expansion $\bar{\alpha}$ and heat capacity $\overline{C_p}$, equation (5) takes a particularly simple form

$$\frac{dR}{dt} = \frac{\bar{\alpha}}{\rho_s\overline{C_p}}\left(K\frac{\partial T}{\partial r}\bigg|_R + \frac{1}{R^2}\right)\int_0^R Ar^2dr \qquad (6)$$

The change of the outer radius of a spherically symmetric body having a constant thermal expansion and heat capacity is proportional to the difference between the rate at which heat is being produced within the volume of the material (the second term on the right hand side of equation (6)) and the rate at which heat is lost at the surface of the spherical body (the first term in equation (6)). The constant of proportionality involves the thermal expansion and the reciprocal of the heat capacity per unit volume of the material. For silicates, the constant of proportionality would be of the order of magnitude

$$\frac{\bar{\alpha}}{\rho_s\overline{C_p}} = \frac{2\times10^{-5}}{3{\cdot}0\times1{\cdot}25\times10^{7}} = 5{\cdot}3\times10^{-13}\ \text{cm}^3\ \text{ergs}^{-1} \qquad (7)$$

Equation (6) states that if the heat flow at the surface is equal to the rate at which heat is produced within the body, then the radius of the body remains constant. A change in radius can come about only through an excess of heat production or an excess of heat loss.

A number of factors are important in considering the balance between heat production and heat loss. The distribution of radioactivity and the effective conductivity are both essential in determining the approach to thermal equilibrium. The greater the near-surface concentration of radioactivity, the shorter is the time required to attain an equality of heat production and heat loss. A high effective conductivity allows more deeply buried radioactivity to come into equilibrium with a near-surface temperature distribution. Of considerable importance is any initial heat that the body may have. This initial heat does not appear explicitly in equation (6) but does contribute to the surface heat flow. A high initial temperature favors the loss of heat and results in a gradually decreasing radius. Further complications are introduced by the fact that the rate of heat production in material having a chondritic composition is at present less by a factor of 8 than the heat production in chondritic materials $4{\cdot}5\times10^9$ years ago. The higher rate of heat production in the past and the high thermal inertia of low-conductivity silicate materials suggest that the higher rate of radioactive decay in the past may play an important part in the thermal and stress history of a planet.

For a numerical example we consider a moon in which the radioactivity is uniformly distributed. Equation (6) then takes the form

$$\frac{dR}{dt} = \frac{\bar{\alpha}}{\rho_s\overline{C_p}}\left(K\frac{\partial T}{\partial r}\bigg|_R + \frac{AR}{3}\right) \qquad (8)$$

MacDonald[4] calculated the thermal history of a Moon with uniform radioactivity having an initial temperature of 600°C. The surface heat flow at present for such a case is 12·4 ergs/cm² sec. The rate of heat production in chondritic material is 1·6 ergs/g year. Substituting these numerical values into equation

(8) and using a coefficient of thermal expansion of 2×10^{-5} °C^{-1} we have that

$$\frac{dR}{dt} = 5{\cdot}3 \times 10^{-13}(-3{\cdot}8 \times 10^{14} + 2{\cdot}8 \times 10^{14}) =$$

$$= -53 \text{ cm}/10^6 \text{ years} \qquad (9)$$

In this case the outer radius is decreasing at the rate of about half a meter per million years. Even though the radioactivity is distributed uniformly through the body, the higher rate of heat production in the past results in a presently greater heat loss than heat production. The radius of the Moon would be decreasing but at a very small rate.

DETAILED CALCULATIONS

The thermal histories of a number of models of the Moon have been calculated, and the corresponding rates of change of the radii have been determined. In going from the thermal history to the rate of change of radius the important parameters are the thermal expansion and effective conductivity. The thermal expansion has been arbitrarily set at 1×10^{-5} °C^{-1}. This value is likely to be low, perhaps by a factor of 2 or 3, if values obtained in the laboratory and values estimated for materials within the earth are at all representative. The rate of change of radius depends linearly on the thermal expansion, and doubling of the mean thermal expansion doubles both the rate of change of the radius and the total change of radius for a fixed interval of time. The thermal conductivity enters in a more complicated way. It is assumed that radiation can contribute to the effective conductivity, and the result of choosing different opacities for the lunar materials is considered in the calculations[4].

Fig. 1 shows the radius and rate of change of radius of the Moon as a function of time, provided the Moon was initially at 0°C throughout. The opacity of the lunar material is taken to be uniform and to be 100 cm^{-1}. The heat capacity is 1·3 J/g °C. The ordinary lattice conductivity is taken as 0·025 J/g sec °C. For such a model the present heat flow after $4{\cdot}5 \times 10^9$ years is 10·3 ergs/cm^2 sec, and the interior temperature is 1780°C[4]. During the

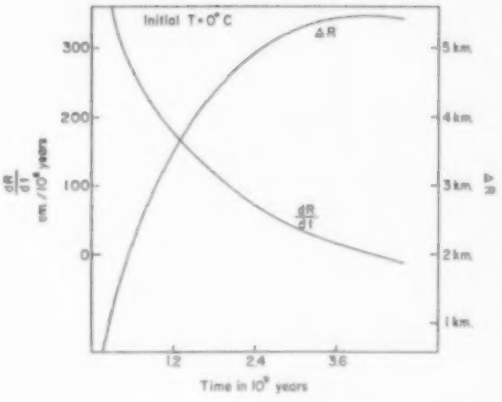


Fig. 1. Change of radius of a Moon initially at 0°C 4·5 x 10^9 years ago. Opacity is 100 cm^{-1}; ordinary lattice conductivity is 0·025 J/g sec °C. Final central temperature and surface heat flow are 1780°C and 10·3 ergs/cm^2 sec, respectively.

first 2×10^9 years, the radius increases by 4·6 km. After that time the rate of increase of radius decreases, with the radius reaching a maximum at $4{\cdot}2 \times 10^9$ years after initial formation. At the present the radius would be decreasing but at a small rate. The striking feature of the curve is the near constancy of the radius over the past $1{\cdot}5 \times 10^9$ years. Over this interval of time the radius has not changed by more than a few hundred meters, while over the past 10^9 years the radius has remained sensibly constant. The sharp initial increase in radius is due to the higher rate of heat production in the past coupled with the low conductivity and the fact that the radioactivity is assumed to be deeply buried.

The case of an initially warm Moon is shown in Fig. 2. The opacity is taken to be 1000 cm^{-1} throughout the body, and therefore the contribution of radiation to energy transport is negligibly small. The initial temperature is taken to be 600°C. The final heat flow is 12·4 ergs/cm^2 sec, while the deep interior temperature is 2380°C. The general features of the variation of radius with time are the same as those for a moon with an initial zero temperature. There is a sharp increase in radius during the early part of the lunar history and a

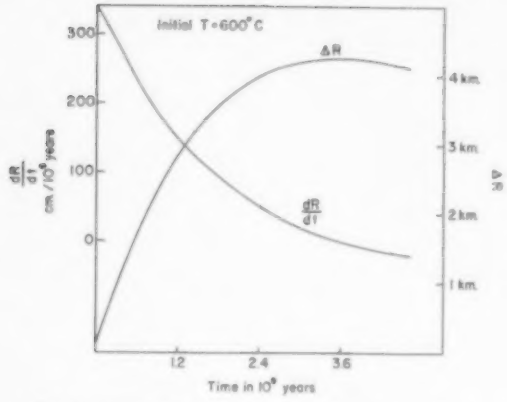


Fig. 2. Change of radius of a Moon initially at 600°C. Opacity is 1000 cm^{-1}. Final central temperature and surface heat flow are 2380°C and 12·4 ergs/cm^2 sec, respectively.

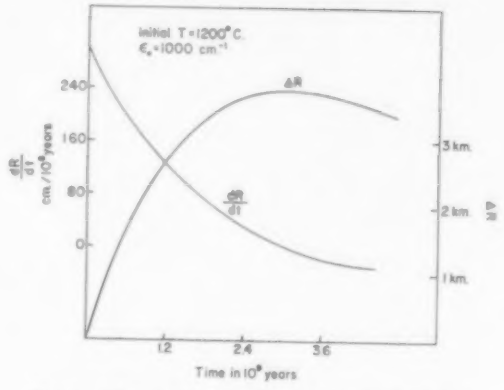


Fig. 3. Change of radius of a Moon initially at 1200°C. Opacity is 1000 cm^{-1}. Final central temperature is 2980°C.

gradual decrease in radius over the past few hundred million years. The total change in radius over the past 10^9 years is only 100 m. The total change in radius over the entire history is less than that in the case of an initially cold Moon. This is entirely due to the higher initial temperature. The heat associated with the higher initial temperature increases the surface heat flow while the rate of heat production remains constant. As in the case of a cold Moon, the radius remains sensibly constant over the past two billion years.

Fig. 3 shows the radial history of a Moon initially at 1200°C and with opacity of 1000 cm^{-1}. The temperature at the center of the moon in this model is 2980°C, far above the melting point of any likely silicate or metallic phase. The model is therefore not appropriate but does indicate the effect on the radius of increasing the initial temperature. With a higher initial temperature, the total increase in radius is now only 3·8 km with a maximum being reached around $3{\cdot}0 \times 10^9$ years after formation of the Moon. Since that time the radius has been decreasing with a total decrease of some 260 m in the last 10^9 years. Fig. 4 illustrates the radial history of a Moon in which the initial temperature is again 1200°C but the opacity is taken as 10 cm^{-1}. In this case the

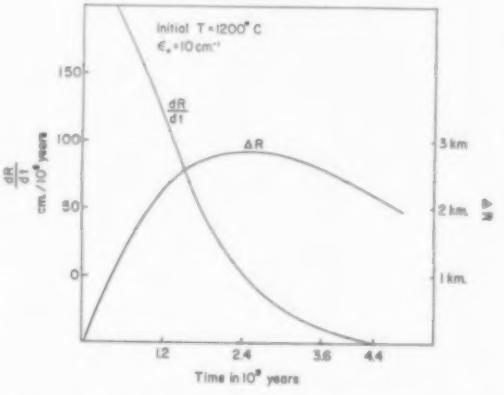


Fig. 4. Change of radius of a Moon initially at 1200°C. Opacity is 10 cm^{-1}. Final central temperature is 2500°C.

final central temperature is 2500°C. The radius increases sharply, as in the other model, during the first 2×10^9 years, reaching a maximum about $2{\cdot}5 \times 10^9$ years after initial formation. Since that time the radius decreases at a rate less than the rate of increase. Over the past 10^9 years, the radius would decrease some 450 m. The maximum increase in radius is 2·8 km. The effects of the high initial temperature and of the contribution of radiation to the transfer of energy are both in the same direction of lowering the total change of radius. Both of these effects tend to increase the surface heat

flow, while the rate of heat production remains constant.

The general features of all four models are the same. The amplitude of the variation in radius depends on the initial temperature and on the mechanism of heat transport. If the heat is transported by radiation or by any other mechanism in which the heat is more efficiently transferred than by ordinary conduction, then the rate of change of radius is less over any interval of time, and the total change in radius over a period of time is proportionally less.

The effect of varying distributions of radioactivity has not been explicitly considered. A near-surface concentration of radioactivity would tend to increase the surface heat flow, while the total heat production would remain constant. As a result, the corresponding changes in radius would be lessened.

The constancy of the Moon's radius can be explained by one of four hypotheses.

(1) The radioactivity of the Moon is far less than that of chondritic materials. In this case the rate of heat production would be sufficiently low so that both terms on the right-hand side of equation (6) would be small.

(2) Radioactivity is more or less uniformly distributed within the Moon, but heat transported by convection is sufficiently important so that the heat flux approximately equals the rate of heat production. Convective heat transport would require molten or perhaps near-molten material and thus would imply a moderate to high initial temperature if the Moon has chondritic composition. An initial temperature of 600°C would probably be sufficient to result in the melting of a portion of the lunar interior. But if melting were important, one would expect evidence for strong surface volcanic activity. Opinions of various observers differ drastically on this point. Urey and Gold consider that the surface has been molded principally by the impact of meteors. However, others would argue for a volcanic origin of many of the lunar features.

(3) The radioactivity of the Moon could be deeply buried and the Moon could have a relatively low initial temperature on the order of several hundred degrees. In this case, as is illustrated in Figs. 1 and 2, the radius of the Moon would remain constant over the past two billion years. This would imply that the lunar surface features are at least 2×10^9 years in age.

(4) The Moon is a chemically differentiated body with the radioactivity concentrated near the surface. In this case the temperature at depth might be well below the melting temperature at all points. Furthermore, the rate of heat flux would equal the rate of heat production. Such a model for the thermal structure of the Moon could well explain the lack of both large-scale faulting and large-scale volcanic activity on the Moon. The principal difficulty in such a model is providing a history for the differentiation of the lunar material.

STRESS DIFFERENCES RESULTING FROM CHANGING TEMPERATURE DISTRIBUTION

Let p_{rr}, $p_{\theta\theta}$ be the radial and tangential stresses arising from an increase in temperature ΔT. The difference between p_{rr} and $p_{\theta\theta}$ is usually used in criteria for determining failure of the material. When this difference exceeds some critical value, the material will undergo fracture or rupture. Hales[5] shows that the rate of increase of maximum stress difference at any radius is given by

$$\left.\frac{d(p_{rr}-p_{\theta\theta})}{dt}\right|_r = \frac{2\mu(1+\sigma)\alpha}{3(1-\sigma)}\left(\left.\frac{dT}{dt}\right|_r - \frac{3}{r^3}\int_0^r r^2 \left.\frac{dT}{dt}\right|_r dr\right) \qquad (10)$$

where μ is rigidity, σ is Poisson's ratio, α the coefficient of thermal expansion. For silicates the constant $\frac{2\mu(1+\sigma)\alpha}{3(1-\sigma)}$ is on the order of about 10 bars per degree. Equation (10) is approximate, since it is assumed that the elastic constants and the thermal expansion remain constant within the Moon.

During the heating-up phase, the tangential stress exceeds the radial stress, and the rate of increase of maximum stress difference is large and negative. At a time when the radius is a

maximum the rate of change of the maximum stress difference vanishes at the surface and during the decrease of the radius of the Moon, the upper parts of the Moon are in compression with the radial stress exceeding the tangential stress. The level at which the rates of increase of tangential and radial stresses are equal depends on time elapsed since the Moon's radius was a maximum. Thus for the case where the Moon is initially at 600°C, the rates of increase of tangential and radial stresses are equal at a depth of 100 km at the present time. For the case in which the Moon was initially at 1200°C, the rates of increase of principal stresses are equal at the present time at a depth of 650 km.

In all the models studied, the average rate of increase of the principal stresses is $|d/dt\,(p_{rr}-p_{\theta\theta})| \geqslant 10$ bars/10^6 years. Using this figure we can estimate the average rate of release of strain energy, provided some assumption is made as to the strength of the lunar materials. For a spherically symmetric distribution the strain energy per unit volume associated with distortion is given by

$$(p_{rr}-p_{\theta\theta})^2/6\mu \qquad (11)$$

The rate of increase of strain energy will be a maximum when the stress differences are on the order of the strength. For a strength of 100 bars, the maximum strain energy per unit volume is on the order of $1{\cdot}4\times10^3$ ergs/cm^3. If the rate of build-up of the maximum stress difference is on the order of 10 bars/1,000,000 years, the rate of release of strain energy by failure at 100 bars maximum stress difference will be on the order of 4×10^{24} ergs/year. At present the Moon should have a high degree of seismic activity unless the radioactivity is concentrated in a thin layer near the surface or unless the radioactivity of the Moon is far less than that of chondritic materials.

The degree of seismic activity varies, of course, among the different models. The model in which there would be the least rate of release of strain energy is that in which the Moon was initially at 1200°C and in which radiation transports a major part of the energy. The highest degree of seismic activity should be expected for a Moon initially cold and heating up with radioactivity distributed uniformly throughout. In the models studied the greatest rate of release of strain energy would be at depths of 100–700 km. The assumption of deeply buried radioactivity implies deep foci for lunar seismic disturbances.

REFERENCES

1. R. B. Baldwin, *The Face of the Moon*. Chicago University Press (1949).
2. H. C. Urey, *The Planets*. Yale University Press, New Haven, Conn. (1951).
3. W. Q. Kennedy, *Quart. J. Geol. Soc., Lond.* **102**, 41 (1946).
4. G. J. F. MacDonald, *J. Geophys. Res.* **64**, 1967 (1959).
5. A. L. Hales, *Mon. Not. R. Astr. Soc., Geophys. Suppl.* **6**, 460 (1953).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 256–260. Printed in Great Britain.

A GEOLOGICAL INTERPRETATION OF THE LUNAR SURFACE

S. MIYAMOTO

Kwasan Observatory, University of Kyoto

(*Received* 29 *December* 1959)

Abstract—This interpretation of lunar surface features is based on the differentiation of siliceous and basaltic substances similar to those of the earth. Our hypothesis is that lunar maria are composed of basaltic matter similar to that of terrestrial ocean basins, and that the lunar continents (bright regions) are composed of siliceous blocks.

1. INTRODUCTION

The Moon's surface is characterized by bright areas covered by ringed structures and great expanses of dark plains. There are many hypotheses dealing with the origin of these surface features. The meteorite theory, first propounded by F. Gruithuisen and recently greatly extended by R. B. Baldwin[1], ascribed the origin of the so-called lunar craters to the bombardment of meteorites. Baldwin, especially, considers the larger maria to be craters produced by the collision of gigantic meteorites. He showed that the kinetic energy of impact is sufficient to produce a melting and spreading of the meteorite, resulting in the formation of maria. A possible argument against the impact hypothesis is the fact that the distribution of the craters is far from random.

The theory of the volcanic origin of the craters is supported by many authors. According to this theory, lunar craters and walled plains correspond to terrestrial volcanoes or calderas. This theory holds that the maria were produced by the immense outflow of lava, which flooded and destroyed pre-existing craters, and which later solidified, forming a nearly level surface as it cooled[2, 3].

Some authors maintain that the impact of meteorites triggered volcanic activity[3] especially for some kinds of explosion craters. While the volcanic theory seems most promising, it is at present far from satisfactory with respect to the interpretation of the great varieties of lunar surface features.

In this paper, a geological interpretation of the surface features is attempted based on the idea that the differentiation of siliceous and basaltic substances took place in the lunar crustal zone as on our earth, and that the lunar maria are composed of basaltic matter corresponding to our ocean basins, while the bright regions are siliceous and correspond to our continental masses. This idea is not new; in fact, R. Henseling[4] put forth this hypothesis in 1927. However, a recent trend holds that the dark basaltic lava outflowed over the underlying light-coloured crust to form the maria, as is suggested by the presence of bright craterlets and isolated mountains in the maria. J. E. Spurr[2], in an important geological investigation, assumed the differentiation of the silicic crust from the basic layer. He identified the former with the ancient crust, which he supposed to cover the entire surface. He suggested that the maria were produced as the result of the collapse of a giant dome of original crust after the release of inflating gases. The subsidence of the crust and the flood of basic magma over it formed the dark expanse of the maria.

2. THE ORIGIN OF MARIA AND CONTINENTS

Near the surface of our earth, the bright-coloured, light-weight siliceous mass differentiates from and floats above the dark and heavy basic mass. The latter forms the ocean basins and the former the continental mass blocks. Although the earth's crust appears to be rigid, it behaves as though it were plastic in the geological time scale and in masses of continental dimension. According to the principle of isostasy, the continental column is thick

beneath high mountain ranges, although the thickness is estimated to be not more than a few tens of kilometers. Under the basaltic layer, the andesitic mantle is believed to extend to great depth. Lava wells up from these layers through fracture systems. Earthquake records indicate that andesite occupies the deeper layers and basalt the shallower.

An inspection of the lunar surface suggests the same differentiation of silicic and basaltic masses as in our earth. However, owing to the weakness of gravitation and rapid cooling, the differentiation may not be so complete as in our earth. In fact, a great expanse of half tone area seen in the other side photograph taken by the Soviet Lunik III suggests incomplete differentiation.

When the continental silicic mass is not thick, a bright silicic film may brush over the basic dark layer in some areas; in other areas the silicic mass may be dotted with configurations resembling isolated icebergs or a chain of icebergs blown together by the convection current in the basaltic magma. Meanwhile, the weak point of the continental crust may under tension be pulled apart by the underlying dark basic magma which is exposed to form a dark-floored, walled plain, or a smaller mare with a circular boundary, such as Mare Crisium.

Among the most impressive features in the maria are the so-called ghost craters, or buried craters, and craters with partly destroyed walls. These objects are often regarded as strong support for the theory that maria were formed by the great inundation of molten lava following subsidence. However, close examination of the lunar surface shows that there are great varieties of ghost craters. Some of them have light-coloured, sharp walls, while others have dark, low walls. In terms of the differentiation hypothesis, large ghost craters with low, dark walls were formed by the fluidal basic crust; the brighter objects are mostly produced of isolated silicic blocks, which may often be shaped complete or incomplete doughnuts affected by local convection pattern.

In and around Mare Nubium, there are abundant ruined rings and ghost craters, strongly suggesting that these areas were once invaded by a lava flow. Palus Epidemiarum is separated westwards from Mare Nubium by the long rampart extending from Cichus northward through Mercator and Campanus to Hippalus. The only possible strait of lava flow from outside is that from the southwestern border of Mare Humorum. But this strait is brighter in colour than the Palus and it is difficult to believe that this was once the actual path of dark lava flow. On the other hand, no sign of a lava source is seen inside the Palus. The floor of this Palus is rather smooth and no wrinkled ridges suggesting the direction of lava flow are seen. Capuanus, situated on the southern shore of this Palus, has a dark floor. Its darkness is comparable to the base of the Palus. But the wall of this crater is complete, except for craterlets on the north, so that it is unlikely that molten lava intruded here. Throughout the lunar surface, most craters with dark floors are located along the shore lines. These facts seem difficult to reconcile with the lava inundation theory.

Palus Epidemiarum provides another interesting topic. A long cleft runs from the north wall of Capuanus westward (Neison's Capuanus rill). It disappears when it meets the mountain range connecting Cichus with Mercator, but reappears in Mare Nubium and eventually reaches Hesiodus. The cleft might have originated when the basaltic basin had been in a plastic state, but the solid silicic mountain range, floating on the basin, remained unaffected. The behaviour of this cleft seems to favour the differentiation hypothesis. While our hypothesis is that the lunar maria are basaltic and correspond to our ocean basins, this does not imply that maria were once covered with water: they may have been waterless oceans ever since the birth of the moon.

3. WALLED PLAINS AND CRATERS

As has been emphasized by J. E. Spurr, lunar geology is fundamentally different from ours. In addition to differences such as the smallness of surface gravity and the extreme tenuity of the atmosphere (even in the earliest lunar

period), the Moon's surface features were formed during the course of the solidification of her crust. On our earth, however, the earliest features have been lost in the geological past by the ceaseless action of erosion and violent crustal movements. The terrestrial volcanic activity of today is the outflow of magma through the fissures of solid crust, and in this respect, our volcanoes are by no means the exact counterpart of lunar craters.

Plato is a walled plain with a dark, smooth floor. It is a giant crater, but a mere dwarf among maria. In fact, its dimensions are much smaller than circular maria such as Mare Crisium and Mare Humorum. In view of our hypothesis, Plato should be grouped with maria because of its dark basaltic floor. When the lunar surface was still in a semi-molten state, a convection cell in the underlying basaltic layer might have broken through the overlying silicic crust and exposed itself. Such an event would take place where the crust is not so thick, hence, near the shore lines.

The circular form and lofty cliffs along the shore line of Mare Crisium may be accounted for by the same process, provided the silicic layer was moderately thick in that area. On the other hand, Mare Humorum, which has no scarps, may be an example of a formation in thin silicic crust.

Goclenius, Magelhaens on the eastern border and many other objects on the northern border of Mare Foecunditatis are smaller than Plato, but they also show dark floors under a high sun. The continental crust is not thick in this region, and the basaltic layer might have been more frequently exposed. Besides the circular formation mentioned above, we find dark but irregular patches in the same area. This area may be called "marshy".

Walled plains and many other lunar craters are by no means explosive in origin. In fact, there are no explosions where the crust is still in a fluidal condition. A violent explosion takes place only after the crust has solidified and enormous pressure has accumulated within it.

At the earliest stage of crustal solidification, local irregularities in the cooling process might have taken place. As is well known, the melting point of magma is greatly reduced by its water content. A spot with a highly volatile constituent would have remained in a fluidal condition down to a comparatively later epoch and, after subsidence, finally fossilized as vertical cavities or craterpits. The importance of the role of a volatile constituent in crater formation should be emphasized in view of the rich cosmic abundance of light elements. We find typical pits of this type in the continent extending to the south-western quarter of the Moon. Craterpits thus formed would naturally take a circular form but they may often be affected by the mesh of cleft systems in the surrounding crust and might take a polygonal appearance. Polygonal craters are so common in lunar surface that some authors choose to speak of them as graben craters.

The terrestrial counterpart of lunar craters may thus be sought for in calderas with the character of graben rather than volcanoes, namely, calderas with very large conduits, comparable in dimension to the outer wall. On the Moon, these collapse calderas with very large conduits seem common, but those like Wargentin with elevated floors are exceptional. This suggests again that the lunar craters were formed before the crust is completely solidified.

Walled plains with light-coloured floors, such as Clavius, may belong to collapse caldera produced in the thick continental crust. Complex plaits or terraces seen on the inner wall of some craters may be the foot prints of successive steps of floor depressions. It must be noted, however, that lunar craters are not necessarily characterized by central mounds or craterlets. They are of secondary importance. Even for those which have central volcanoes, these volcanoes do not supply the generating power of crater formation.

Contrary to the earth's surface, craters, not mountains, constitute the dominant feature of the lunar landscape. While there are mountains—the lofty mountain range surrounding Mare Crisium, the Apennines bordering Mare

Imbrium, Sovetsky mountain range — there seems to be no real orogeny on the moon.

The Taurus mountains, west of Mare Serenitatis, are quite different from the Apennines, presenting to the viewer the appearance of a slushy road. The mottling of semi-molten craters may be due to the brindled cooling of the crust. The highland extending south of Mare Crisium shows a similar landscape. In this region, depressions are somewhat regularly arranged with superposed chains of calderas. When a chain of calderas is isolated, it may appear to be a valley, like Rheita.

4. THE EXPLOSIVE NATURE OF SILICIC MAGMA

It is well known that explosive activity is more common in volcanoes that eject silicic lava than in basaltic volcanoes. The explosive power of magma depends upon how much gas it contains. Lunar continents are crowded with innumerable craters and the volatile constituents of silicic magma must have played an important role in their formation.

Basic magma contains lesser amounts of gases. It is denser and less viscous so that the condition in the maria is not so favourable for crater formation. Some of the so-called ghost craters with dark-colored, low walls may be counterparts of continental craters. The oldest craters, formed when the basin was still in a plastic state, are of large dimension and the surrounding walls are low. An interesting example is the large ghost crater situated southwest of Arago at the centre of Mare Tranquilitatis. Its shadowy walls are almost indistinguishable from the neighbouring wrinkled ridges. Later products, formed after the basin acquired hardness, have reduced the dimensions of the crater and, at the same time the surrounding walls have become more conspicuous.

The domes that abound in maria suggest the evolution of craterlets in the maria. Close examination reveals that most of these domes are equipped with a central pit. Giant domes near Arago in Mare Tranquilitatis have small pits, but those in Mare Nubium have pits that are comparatively larger and are easily detectable. These domes are seldom associated with fissure systems. Domes arranged in a sequence of increasing pit size appear to be genetically related to craterlets and suggest collapse caldera as the nature of isolated craterlets distributed at random in the maria.

Let us now turn to the bright craterlets. When mass blocks of high silicic content were spotted amid the ocean basin, such spots would often be unstable, leading to crater formation, because of high contents of volatile constituents. Bright craterlets observed in the maria may have been formed in such a way. Another possibility is the violent outflow of deeper andesitic magma in a later stage, after the solidification of the ocean basin.

Of course, we do not reject entirely the idea of the inundation of basaltic lava over the pre-existing craters in some areas. In fact, the anonymous crater due west of Delambre on the eastern border of Mare Tranquilitatis may be an example of this. However, one point must be raised against the lava flow hypothesis. A great lava flow would surely destroy many craterlets and might bury them entirely, leaving no trace of their existence. But even if we take this into account, present maria possess too few craters. Mare Nubium, for example, is covered with a comparatively thin lava layer; buried craters and irregular patches of brightness are seen clearly. But the number of such craters is few in contrast with the great crowd of craters in the adjacent continent extending southwards.

In summary, lunar craters may be classified in a basic two-dimensional array. The first co-ordinate is the chemical composition of the crust, namely, the relative abundance of silicates. The second co-ordinate is the plasticity of the crust or the epoch in the cooling period. Of course, metamorphism cannot be included in this scheme.

5. RING PLAIN AND RAY SYSTEM

After the advanced cooling of the lunar surface and solidification of the crust, the real volcanoes might have come into being. We see what appears to be lava flows and ejectamenta around Copernicus, Aristillus, etc., which are

characteristic of objects of this type but absent in other caldera-type craters.

A full Moon photograph shows that lunar continents are dotted with bright craters. Most of them are small, and giants such as Tycho, Copernicus and Kepler are few in number. Under a high sun, they shine brilliantly, often accompanied by a mysterious ray system. The continental corridor surrounded by Mare Imbrium, Oceanus Procellarum and Sinus Roris is most spectacular. At full Moon, this area is dotted with numerous "stars", contrasting sharply with the surrounding darkness of the maria. After D. Alter, they are called explosion craters. They may be ascribed to the explosive overflow of silicic magma, rich in gaseous constituents, through fissures or weaknesses in the crust. The spectacular ray system radiating from the explosion crater is usually considered as emanating from the deposit of gases and cinders exhaled from volcanoes. Such explosions might have been the final activity of the dying moon.

In comparing Copernicus with Aristillus, we find that the latter is darker and has no ray system, but both show ejectamenta around them. The brightness difference may be attributed to the difference in the chemical composition of the lava. The Aristillus lava may be more basic and hence less explosive. There is apparently less lunar orogeny and hence the supply of heat maintaining the volcanic activity must have declined rapidly. The lunar landscape is in fact the product of the cooling period.

Acknowledgements—The author wishes to thank Dr. A. G. Wilson for his encouragement and valuable comments. He is also grateful for the kind cooperation of Mr. M. Matsui, of the Kwasan Observatory staff, who secured many excellent photographs of the moon. This work is based entirely on his material.

REFERENCES

1. R. B. Baldwin, *The Face of the Moon*. University of Chicago Press (1949).
2. J. E. Spurr, *Geology Applied to Selenography*. Science Press, Lancaster, Pa. (1944).
3. H. P. Wilkins and P. Moor, *The Moon*. Macmillan, London (1954).
4. R. Henseling, *Der Mond und die Planeten*. Stuttgart (1927).

Planet. Space Sci. Pergamon Press 1960. Vol. 2, pp. 261–262. Printed in Great Britain.

RESEARCH NOTE

Some comments on the penetration of whistlers through ionospheric layers

(*Received* 22 *January* 1960)

A paper[1] recently published in this journal contains some characteristic properties of whistlers. It was found that whistlers appeared in groups during periods of half an hour up to several hours—whistler situations. No whistlers occur between these whistler situations. As a rule in such situations, the number of whistlers generally amounts to three during observation periods of 5 min. On a day with intense activity, characterized as a "whistler storm", groups of about twenty whistlers can appear during subsequent periods of 30 min. During periods between whistler situations total cessation of whistlers occurred, sometimes extending over several days. The cause of these cessations could possibly be that no thunderstorms of sufficient intensity occurred within the observation area. This was however not the case, as the mean level of the field strength values of lightning discharges in the area was at least as high as during active whistler situations.

Another important fact was that a thunderstorm area located by a direction finder in a certain direction was accompanied by whistlers while a simultaneous thunderstorm area at about the same distance but in quite another direction was not accompanied by whistlers. It was also found that whistlers from a thunderstorm in a certain direction could cease while the thunderstorm was still going on. On the other hand, a whistler activity could start suddenly in a thunderstorm occurring in quite a different direction. These singularities of the appearance of whistlers observed during the thunderstorm season of 1958 have been fully confirmed during more extended not yet published investigations during the season of 1959.

Referring to these facts it seems necessary to present, if possible, the prevailing causes why whistlers are not generated in all thunderstorm situations. From our general experience of the mechanism of lightning discharges it is concluded that in all the thunderstorm situations mentioned a considerable number of lightning flashes with the same qualities has been developed. Among these flashes there must exist discharge forms well fitted for the generation of low-frequency wave-packets around 5 kc which is especially favourable for the development of whistlers. When, in spite of these conditions, whistlers did not develop there must be some specific source of hindrance not within the proper thunderstorm atmosphere but located considerably higher up above the tops of the thunder-clouds in the ionospheric regions.

Soon after our paper[1] had been issued, K. G. Budden[2] published a theoretical analysis which may explain why whistlers cannot always penetrate the ionospheric layers. He applied the magneto-ionic theory on such irregularities in the ionosphere as are small compared with the wave-lengths of whistlers. The irregularities in one dimension were treated by statistical methods and the irregularities in three dimensions by successive approximations.

Budden's conclusive remarks were that irregularities which in their extension are small compared with the wave-lengths of whistlers may play an important part and prevent the propagation of very low-frequency radio waves through the ionosphere.

The next radical research step would obviously be to investigate the mentioned irregularities experimentally. In this respect a very comprehensive recent publication[3] dealing with the stratification in the lower ionosphere is of special interest. The authors recommended electron profile measurements by rocket flights for detailed investigations of the ionosphere. This method might result in some practical information of the existence of the theoretically treated small irregularities in the ionospheric layers.

H. NORINDER

University of Uppsala

REFERENCES

1. H. Norinder and E. Knudsen, *Planet. Space Sci.* **1**, 173 (1959).
2. K. G. Budden, *J. Res. Nat. Bur. Stand.* **63D**, 135 (1959).
3. C. Ellyett and J. M. Watts, *J. Res. Nat. Bur. Stand.* **63D**, 117 (1959).

ERRATUM

S. F. SINGER, Properties of the Upper Atmosphere and Their Relation to the Radiation Belts of the Earth, *Planet. Space Sci.* **2**, Nos. 2/3, 165 (1960).

An error occurring in the calculation of the concentration of hydrogen atoms has been corrected as follows. In Fig. 1, replace the values appearing there by the following Table, which is taken from a paper by E. J. Öpik and S. F. Singer, *Physics of Fluids*, 1960, to be published.*

Density Distribution of Hydrogen in the Earth's Exosphere for Various Temperatures of Space

r/R	$T=$ 1250°	$T=$ 1500°	$T=$ 1800°	$T=$ 3000°
1·000	100·00	100·00	100·00	100·00
1·111	54·9	58·95	62·2	67·3
1·250	30·3	34·75	38·7	45·9
2·000	4·1	5·708	7·5	12·4
3·333	0·96	1·108	2·2	4·4
5·000	0·030	0·347	0·096	0·252

Note also that recent results by Tousey's group at the Naval Research Laboratory have given a normalization for the atomic hydrogen distribution which fixes the concentration at the base of the exosphere at about 10^4 atoms cm^3.